

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:33:06 ; Search time 83 seconds
(without alignments)
47.809 Million cell updates/sec

Title: SEQID2-NAT3

Perfect score: 136
Sequence: 1 RPNYSRRPKGVKHLKDFPILPGEI 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_l9Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	93.4	56	ABJ04946	A2 domain of fVII
2	127	93.4	294	AAG79422	Human factor VIII
3	127	93.4	368	AAR73020	Human Factor-VIII
4	127	93.4	720	AAR74088	Factor-VIII heavy
5	127	93.4	729	AAR74089	Factor-VIII heavy
6	127	93.4	740	AAR74090	Human Factor-VIII
7	127	93.4	740	AAR74091	Human Factor-VIII
8	127	93.4	740	AAR74092	Human Factor-VIII
9	127	93.4	740	AAR74093	Factor-VIII heavy

10	127	93.4	740	16	AAR76982	Human factor VIII
11	127	93.4	1383	18	AAW33227	Procoagulant-activ
12	127	93.4	1383	18	AAW33228	Procoagulant-activ
13	127	93.4	1383	18	AAW33229	Procoagulant-activ
14	127	93.4	1424	9	AAP80268	Modified factor VI
15	127	93.4	1424	10	AAP91169	Sequence of 740 Ar
16	127	93.4	1424	22	AAW48842	Mutant mature huma
17	127	93.4	1424	23	AAO18622	Human mature B-dom
18	127	93.4	1425	9	AAP80267	Modified factor VI
19	127	93.4	1438	21	AAW01262	B-domain deleted f
20	127	93.4	1440	12	AAW12971	Factor VIII:SQ. U
21	127	93.4	1445	23	ABG92540	LE B-domain-dele
22	127	93.4	1447	23	ABG92541	Sarg B-domain-dele
23	127	93.4	1457	19	AAW46246	Human factor VIII
24	127	93.4	1457	19	AAW44372	Human Factor VIII
25	127	93.4	1457	20	AAW21675	Beta-domain delete
26	127	93.4	1459	22	AAE10827	Human factor VIII
27	127	93.4	1459	22	AAE10832	Human factor VIII
28	127	93.4	1459	22	AAE10833	Human factor VIII
29	127	93.4	1471	18	AAW23414	Human B-domain del
30	127	93.4	1471	22	AAW67959	Amino acid sequenc
31	127	93.4	1516	9	AAP80265	Modified factor VI
32	127	93.4	1661	18	AAW18670	Factor VIII-dB695-
33	127	93.4	2098	17	AAR86863	Factor-VIII. Homo
34	127	93.4	2332	8	AAP71726	Factor VIII:c varia
35	127	93.4	2332	8	AAP71727	Factor VIII:c varia
36	127	93.4	2332	8	AAP71728	Factor VIII:c varia
37	127	93.4	2332	8	AAP71729	Factor VIII:c varia
38	127	93.4	2332	14	AAW43257	Human Factor VIII.
39	127	93.4	2332	18	AAW33222	Procoagulant-activ
40	127	93.4	2332	18	AAW33223	Procoagulant-activ
41	127	93.4	2332	18	AAW33224	Procoagulant-activ
42	127	93.4	2332	18	AAW33225	Procoagulant-activ
43	127	93.4	2332	18	AAW33226	Procoagulant-activ
44	127	93.4	2332	19	AAW53483	Human factor VIII.
45	127	93.4	2332	19	AAW44132	Homo sapiens modif

ALIGNMENTS

RESULT 1
ABJ04946
ID ABJ04946 standard; Protein; 56 AA.
AC ABJ04946;
XX AC
XX AC
XX 30-OCT-2002 (first entry)
XX DE
XX A2 domain of fVII epitope SEQ ID No 2.
XX DE
XX Haemostatic; antibody inhibitor; factor VIII; T cell; immune response;
XX KW haemophilia A; acquired haemophilia; human factor VIII.
XX OS
XX Homo sapiens.
XX PN WO200260917-A2.
XX PD 08-AUG-2002.
XX PF 30-NOV-2001; 2001WO-US44945.
XX PR 01-DEC-2000; 2000US-250430P.
XX PA (MINU) UNIV MINNESOTA.
XX PL Conti-fine BM;
XX DR WPI; 2002-627462/67.
XX PT New peptides or their variants, useful for preventing, treating or
XX PT inhibiting aberrant or pathogenic production of antibodies specific for
XX PT factor VIII, particularly useful for treating hemophilia A or acquired

```

CC Fragments of these polypeptides inhibit the interaction of blood
CC coagulation factor VIIIa with blood coagulation factor IXa. They also
CC inhibit the activation of blood coagulation factor X, or inhibit blood
CC coagulation. These peptide fragments are useful in preventing and/or
CC treating coagulation disorders, in particular cardiovascular disorders
CC such as thrombosis, atherosclerosis and restenosis.
XX Sequence 294 AA;
SQ
Query Match 93.4%; Score 127; DB 23; Length 294;
Best Local Similarity 96.0%; Pred. No. 1.9e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 55 RPLYSRRLPKGVKHLKDFPILPGEI 79
II IIIIIIIIIIIIIIIIIIIIII
RESULT 3
AAR73020
ID AAR73020 standard; peptide: 368 AA.
AC AAR73020;
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
DE Human Factor-VIII fragment.
XX
KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
OS Homo sapiens.
XX
PN WC9513301-A1.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK00424.
XX
PR 12-NOV-1993; 93DK-0001281.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Persson E;
XX
WP I: 1995-194038/25.
XX
PT Crosslinked factor VIII polypeptide which is stable - is prepd. using
PT bis(sulphosuccinimidylyl) suberate or disuccinimidylyl suberate in the
PT presence of polysorbate 80 to produce a coagulant with long lasting
PT activity
XX
PS Disclosure: Page 19; 36pp; English.
XX
This is a fragment corresponding to internal AAs 373-740 of human
Factor-VIII which may be crosslinked resulting in
increased stability and retention of high activity over extended
periods of time after activation by thrombin. The polypeptide is
used to prevent or treat diseases caused by the absence or deficiency
of factor-VIII in a subject such as haemophilia.
(C updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 368 AA;
Query Match 93.4%; Score 127; DB 16; Length 368;
Best Local Similarity 96.0%; Pred. No. 2.4e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 112 RPLYSRRLPKGVKHLKDFPILPGEI 136
II IIIIIIIIIIIIIIIIIIIIII

```

```

RESULT 4
AAR74088
ID AAR74088 standard; protein: 720 AA.
XX
XX AC AAR74088;
XX
DT 25-MAR-2003 (updated)
DT 04-NOV-1995 (first entry)
XX
XX Factor-VIII heavy chain N-terminal fragment.
DE
DE human: Factor VIII; heavy chain; N-terminal fragment;
KW thrombin cleavage; blood-clotting.
XX
XX Homo sapiens.
XX
XX WO9513300-A1.
XX
XX 18-MAY-1995.
XX
XX Factor-VIII heavy chain N-terminal fragment.
DE
DE human: Factor VIII; heavy chain; N-terminal fragment;
KW thrombin cleavage; blood-clotting.
XX
XX Homo sapiens.
XX
XX WO9513300-A1.
XX
XX 18-MAY-1995.
XX
XX 10-NOV-1994; 94WO-DK00423.
XX
XX 12-NOV-1993; 93DK-0001280.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Ezban Rasmussen M, Kjalke M;
XX
XX WPI: 1995-194037/25.
XX
XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
XX Al-A2 domain - are easier to produce recombinantly and retain coagulant
XX activity, may be used to treat patients who have developed antibodies to
XX C-terminal epitope(s) of Factor VIII
XX
XX Claim 3; Page 24-26; 51pp: English.
XX
XX The sequence represents N-terminal residues 1-720 of a human Factor-
XX VIII heavy chain. The sequence is shorter than the Al-A2 domain,
XX and is produced by treating a polypeptide containing the full Al-A2
XX domain of full-length Factor-VIII with a protease, e.g. thrombin.
XX The fragment has the same specific activity as full-length
XX Factor-VIII in a chromogenic assay, but activity is a factor of two
XX lower in a clotting assay, and the fragment is activated by thrombin
XX at a slower rate and to a lower level than fragments 1-740 (AAR74090),
XX 1-729 (AAR74089) and plasma Factor-VIII. The fragment may be produced
XX recombinantly to reduce production costs and improve safety, and
XX production levels and stability are higher than for the full-length
XX form. The fragment may be used to treat patients who have developed
XX antibodies against epitopes in the C-terminal part of the heavy chain.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 720 AA;
XX
XX Query Match 93.4%; Score 127; DB 16; Length 720;
XX Best Local Similarity 96.0%; Pred. No. 5e-10;
XX Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPNYSRRLPGVKVHLKDFPILPGEI 25
XX || |||||
XX DB 484 RPLYSRRLPGVKVHLKDFPILPGEI 508
XX
XX RESULT 5
AAR74089
ID AAR74089 standard; protein: 729 AA.
XX
XX AC AAR74089;
XX
XX 25-MAR-2003 (updated)
XX 04-NOV-1995 (first entry)
XX
XX Factor-VIII heavy chain N-terminal fragment.
DE
DE human: Factor VIII; heavy chain; N-terminal fragment;
KW thrombin cleavage; blood-clotting.
XX
XX Homo sapiens.
XX
XX WO9513300-A1.
XX
XX 18-MAY-1995.
XX
XX 10-NOV-1994; 94WO-DK00423.
XX
XX 12-NOV-1993; 93DK-0001280.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Ezban Rasmussen M, Kjalke M;
XX
XX WPI: 1995-194037/25.
XX
XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
XX Al-A2 domain - are easier to produce recombinantly and retain coagulant
XX activity, may be used to treat patients who have developed antibodies to
XX C-terminal epitope(s) of Factor VIII
XX
XX Claim 3; Page 24-26; 51pp: English.
XX
XX The sequence represents N-terminal residues 1-720 of a human Factor-
XX VIII heavy chain. The sequence is shorter than the Al-A2 domain,
XX and is produced by treating a polypeptide containing the full Al-A2
XX domain of full-length Factor-VIII with a protease, e.g. thrombin.
XX The fragment has the same specific activity as full-length
XX Factor-VIII in a chromogenic assay, but activity is a factor of two
XX lower in a clotting assay, and the fragment is activated by thrombin
XX at a slower rate and to a lower level than fragments 1-740 (AAR74090),
XX 1-729 (AAR74089) and plasma Factor-VIII. The fragment may be produced
XX recombinantly to reduce production costs and improve safety, and
XX production levels and stability are higher than for the full-length
XX form. The fragment may be used to treat patients who have developed
XX antibodies against epitopes in the C-terminal part of the heavy chain.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 720 AA;
XX
XX Query Match 93.4%; Score 127; DB 16; Length 720;
XX Best Local Similarity 96.0%; Pred. No. 5e-10;
XX Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPNYSRRLPGVKVHLKDFPILPGEI 25
XX || |||||
XX DB 484 RPLYSRRLPGVKVHLKDFPILPGEI 508
XX
XX RESULT 6
AAR76961
ID AAR76961 standard; protein: 740 AA.
XX
XX AC AAR76961;
XX
XX 09-MAR-1996 (first entry)
XX
XX Human Factor-VIII derivative.
DE
DE Factor-VIII; therapeutic; blood-clotting.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 692
XX /label= absent or Ala, Thr, Ser, Gly or Asp
XX Misc-difference 720
XX /label= absent or Gln, Ser, Thr, Val or Ala
XX Misc-difference 729
XX /label= absent or Val, Ala or Ile
XX

```

```

PN WO9518827-AL.
XX
PD 13-JUL-1995.
XX
PF 06-JAN-1995; 95WO-DK00008.
XX
PR 07-JAN-1994; 94DK-0000032.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX WPI: 1995-255039/33.
XX
PT Novel factor VIII derivative used to treat haemophilia - and
PT comprises a functional A2 domain containing a mutation at one or
PT more Cys residues.
XX
PS Disclosure: Page 14-17; 30pp; English.
XX
PX The new Factor-VIII derivative comprises a functional A2 domain in
CC which Cys-692 is deleted or replaced with another amino acid
CC residue, preferably Ser (see AAR76962). Alternatively, Glu-720 and/or
CC Tyr-729 are deleted or substituted with various amino acids (as in
CC the features). The new derivative has the same activity as the wild-
CC type Factor-VIII but with improved stability (the activity is
CC maintained for a longer period compared to the rapid decline of the
CC activity of wt Factor-VIII). The new derivative can be used in a
CC composition for treating diseases caused by an absence or deficiency
CC of Factor-VIII, especially haemophilia.
XX
SQ Sequence 740 AA:

Query Match 93.4%; Score 127; DB 16; Length 740;
Best Local Similarity 96.0%; Pred. No. 5,2e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 7
AAR76962
ID AAR76962 standard: protein: 740 AA.
XX
AC AAR76962:
XX
DT 09-MAR-1996 (first entry)
XX
DE Human Factor-VIII derivative.
XX
KW Factor-VIII; therapeutic; blood-clotting.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 692 /label= Cys substituted by Ser
FT Misc-difference 720 /label= absent or Gln, Ser, Thr, Val or Ala
FT Misc-difference 729 /label= absent or Val, Ala or Ile
XX
PN WO9518827-AL.
XX
PD 13-JUL-1995.
XX
PF 06-JAN-1995; 95WO-DK00008.
XX
PR 07-JAN-1994; 94DK-0000032.
XX
PA (NOVO ) NOVO-NORDISK AS.

```

```

XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX WPI: 1995-255039/33.
XX
PT Novel factor VIII derivative used to treat haemophilia - and
PT comprises a functional A2 domain containing a mutation at one or
PT more Cys residues.
XX
XX Disclosure: Page 18-20; 30pp; English.
XX
PX The new Factor-VIII derivative comprises a functional A2 domain in
CC which Cys-692 is replaced with Ser. For other (less preferred)
CC substitutions at this site, see AAR76961. Alternatively, Glu-720 and/or
CC Tyr-729 are deleted or substituted with various amino acids (as in
CC the features). The new derivative has the same activity as the wild-
CC type Factor-VIII but with improved stability (the activity is
CC maintained for a longer period compared to the rapid decline of the
CC activity of wt Factor-VIII). The new derivative can be used in a
CC composition for treating diseases caused by an absence or deficiency
CC of Factor-VIII, especially haemophilia.
XX
SQ Sequence 740 AA:

Query Match 93.4%; Score 127; DB 16; Length 740;
Best Local Similarity 96.0%; Pred. No. 5,2e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 8
AAR73021
ID AAR73021 standard: peptide: 740 AA.
XX
AC AAR73021;
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX
DE Human Factor-VIII N-terminal fragment.
XX
KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX
OS Homo sapiens.
XX
PN WO9513301-AL.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK00424.
XX
PR 12-NOV-1993; 93DK-0001281.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Persson E;
XX
XX WPI: 1995-194038/25.
XX
PT Crosslinked Factor VIII polypeptide which is stable - is prepd. using
PT bis(sulphosuccinimidy)l suberate or disuccinimidy)l suberate in the
PT presence of polysorbate 80 to produce a coagulant with long lasting
PT activity
XX
PS Disclosure: Page 21; 36pp; English.
XX
PX This is the N-terminal fragment of human Factor-VIII which may be
CC crosslinked resulting in increased stability and retention of high
CC activity over extended periods of time after activation by thrombin.
CC The polypeptide is used to prevent or treat diseases caused by the

```



```

RESULT 11
AAW33227
ID AAW33227 standard; protein: 1383 AA.
XX
AC AAW33227;
XX
DT 30-APR-1998 (first entry)
XX
DE Procoagulant-active human factor VIII:C (FVIII) mutant protein:.
XX
KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
XX recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..346 /note= "factor VIIIA heavy chain"
FT Region 741..1383
FT Domain 1..329 /note= "factor VIIIA light chain"
FT Domain 1..179 /note= "A1 domain"
FT Domain 187..329 /note= "plastocyanin-like domain 1"
FT Domain 380..711 /note= "plastocyanin-like domain 2"
FT Domain 711..746 /note= "A2 domain"
FT Misc_feature /note= "a spacer of the sequence
FT SFQNSRHPSTRQKQFNATIPENDIEKTDPMF
FT AHRTPMKIQNVSSDLLMLL is inserted
FT between domains A2 and A3"
FT
FT Domain 380..554
FT /note= "plastocyanin-like domain 3"
FT Domain 564..711
FT /note= "plastocyanin-like domain 4"
FT Domain 746..1073
FT /note= "A3 domain"
FT Domain 1073..1221
FT /note= "C1 domain"
FT Domain 1226..1378
FT /note= "C2 domain"
FT Cleavage-site 372..373
FT /note= "by thrombin"
FT Disulfide-bond 153..179
FT /note= "probable"
FT Disulfide-bond 528..554
FT /note= "probable"
FT Misc-difference 740
FT /label= R740A
FT /note= "wild type Arg replaced with Ala"
XX
XX W09740145-Al.
XX
XX 30-OCT-1997.
XX
XX 24-APR-1997; 97WO-US06563.
XX
XX 15-MAY-1996; 96US-0017785.
XX 24-APR-1996; 96US-0016117.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Amano K, Kaufman RJ, Pipe SK;
XX WPI; 1997-535830/49.
XX
XX Modified human pro-coagulant active factor VIII - can be

```

```

PT administered to haemophiliacs, i.e. factor VIII replacement therapy
XX
XX Claim 20: Page -: 57pp; English.
XX
CC The present sequence represents a novel pro-coagulant active factor
CC VIII(FVIII) mutant protein, comprising a deletion of the B domain; and
CC von Willebrand factor binding site, a mutation at Arg740 and an addition
CC of an amino acid sequence spacer between the A2 and A3 domains. Factor
CC VIII, along with calcium and phospholipid, acts as a cofactor for factor
CC IXA, when it converts factor X to the activated form (factor XA). FVIII
CC is the coagulation factor deficient in the X-chromosome-linked bleeding
CC disorder haemophilia A. Several other mutant FVIII proteins have also
CC been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
CC capable of recombinant secretion at higher levels than typically obtained
CC with wild type FVIII and retains pro-coagulant activity. The FVIII
CC mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated
CC protein C (APC) cleavage. The present FVIII mutant can form a more
CC stable configuration, and have an approximate 5-fold increase in
CC specific activity compared to purified wild type FVIII, while increasing
CC their binding affinity to von Willebrand factor improves their
CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
CC FVIII replacement therapy, while the nucleic acid molecule can be used
CC for gene therapy.
CC note: this sequence does not appear in the specification: it was created
CC using sequences from the given references.
XX
XX Sequence 1383 AA;
SQ
Query Match 93.4%; Score 127; DB 18; Length 1383;
Best Local Similarity 96.0%; Pred. No. 1e-09;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RPNVSRRLPGVGHKDFPILPGEI 25
DB 484 RPLYSRRLPGVGHKDFPILPGEI 508
RESULT 12
AAW33228
ID AAW33228 standard; protein: 1383 AA.
XX
AC AAW33228;
XX
DT 30-APR-1998 (first entry)
XX
DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..346 /note= "factor VIIIA heavy chain"
FT Region 741..1383
FT Domain 1..329 /note= "factor VIIIA light chain"
FT Domain 1..179 /note= "A1 domain"
FT Domain 187..329 /note= "plastocyanin-like domain 1"
FT Domain 380..711 /note= "plastocyanin-like domain 2"
FT Domain 711..746 /note= "A2 domain"
FT Misc_feature /note= "a spacer of the sequence
FT SFQNSRHPSTRQKQFNATIPENDIEKTDPMF
FT AHRTPMKIQNVSSDLLMLL is inserted
FT between domains A2 and A3"
FT
FT Domain 380..554
FT /note= "plastocyanin-like domain 3"
FT Domain 564..711
FT /note= "plastocyanin-like domain 4"
FT Domain 746..1073
FT /note= "A3 domain"
FT Domain 1073..1221
FT /note= "C1 domain"
FT Domain 1226..1378
FT /note= "C2 domain"
FT Cleavage-site 372..373
FT /note= "by thrombin"
FT Disulfide-bond 153..179
FT /note= "probable"
FT Disulfide-bond 528..554
FT /note= "probable"
FT Misc-difference 740
FT /label= R740A
FT /note= "wild type Arg replaced with Ala"
XX
XX W09740145-Al.
XX
XX 30-OCT-1997.
XX
XX 24-APR-1997; 97WO-US06563.
XX
XX 15-MAY-1996; 96US-0017785.
XX 24-APR-1996; 96US-0016117.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Amano K, Kaufman RJ, Pipe SK;
XX WPI; 1997-535830/49.
XX
XX Modified human pro-coagulant active factor VIII - can be

```



```

XX 24-APR-1997; 97WO-US06563.
XX
XX
XX 15-MAY-1996; 96US-0017785.
XX 24-APR-1996; 96JS-0016117.
XX
XX (JNMI ) UNIV MICHIGAN.
XX
XX Amano K, Kaufman RJ, Pipe SW;
XX WPI; 1997-535830/49.
XX
XX Modified human pro-coagulant active factor VIII - can be
XX administered to haemophiliacs, i.e. factor VIII replacement therapy
XX
XX Claim 19; Page -: 57pp; English.
XX
XX The present sequence represents a novel pro-coagulant active factor
XX VIII (FVIII) mutant protein, comprising a deletion of the B domain and
XX von Willebrand factor binding site, mutations F109S, R740A and addition
XX of an amino acid sequence spacer between the A2 and A3 domains. Factor
XX VIII, along with calcium and phospholipid, acts as a cofactor for factor
XX IXA, when it converts factor X to the activated form (factor XA). FVIII
XX is the coagulation factor deficient in the X-chromosome-linked bleeding
XX disorder haemophilia A. Several other mutant FVIII proteins have also
XX been created (see AAW3222-29). The FVIII mutant F309S (AAW3225) is
XX capable of recombinant secretion at higher levels than typically obtained
XX with wild type FVIII and retains pro-coagulant activity. The FVIII
XX mutant R336I (AAW3222) and R562K (AAW3223) are resistant to activated
XX protein C (APC) cleavage. The present FVIII mutant can form a more
XX stable configuration, and have an approximate 5-fold increase in
XX specific activity compared to purified wild type FVIII, while increasing
XX their binding affinity to von Willebrand factor improves their
XX stability. The FVIII proteins can be administered to haemophiliacs, i.e.
XX FVIII replacement therapy, while the nucleic acid molecule can be used
XX for gene therapy.
XX note: this sequence does not appear in the specification; it was created
XX using sequences from the given references.
XX
XX Sequence 1383 AA;
XX
XX Query Match 93.4%; Score 127; DB 18; Length 1383;
XX Best Local Similarity 96.0%; Pred. No. 1e-09;
XX Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
XX || ||||| ||||| ||||| ||||| |||||
XX DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508
XX
XX RESULT 14
XX AAP80268
XX ID AAP80268 standard; protein; 1424 AA.
XX
XX AC AAP80268;
XX
XX XX 25-MAR-2003 (updated)
XX DT 10-OCT-1990 (first entry)
XX
XX XX Modified factor VIII:C sequence with the R740-D1658 deletion.
XX
XX XX Modified factor VIII:C; maturation polypeptide; haemophilia;
XX KW blood coagulation; RD deletion; procoagulant.
XX
XX OS Homo sapiens.
XX
XX XX W08800831-A.
XX PN
XX XX 11-FEB-1988.
XX PD
XX XX 31-JUL-1987; 87WO-US01814.
XX PF
XX XX 01-AUG-1986; 86US-0893375.
XX PR

```

```

XX (BIOJ ) BIOGEN NV.
XX PA (PASE/) PASEK M P.
XX
XX PT Pasek MP;
XX
XX WPI; 1988-049866/07.
XX DR N-PSDB; AAN80447.
XX
XX New DNA sequences encoding modified factor VIII:C - with deletion of DNA
XX encoding maturation polypeptide, useful for high yield transformation.
XX
XX Claim 3; Page 60-61-62-63; 97pp; English.
XX
XX The RD deletion removes the DNA from Ser 741 to Ser 1657.
XX A major part of the sequence encoding the maturation polypeptide of
XX factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.
XX The full length Factor VIII:C cDNA has two changes with respect to the
XX published sequence (EPO application 160457):
XX CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
XX (Phe to Leu). The product is produced in approx. 20 times higher
XX yields than previous recombinant produced factor VIII:C and are more
XX easily purified. The peptide is used for treating haemophilia A, both
XX acute and prolonged bleeding.
XX See also AAN80444 and AAN80446.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1424 AA;
XX
XX Query Match 93.4%; Score 127; DB 9; Length 1424;
XX Best Local Similarity 96.0%; Pred. No. 1e-09;
XX Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
XX || ||||| ||||| ||||| ||||| |||||
XX DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508
XX
XX RESULT 15
XX AAP91169
XX ID AAP91169 standard; protein; 1424 AA.
XX
XX AC AAP91169;
XX
XX DT 25-MAR-2003 (updated)
XX DT 26-JUN-1990 (first entry)
XX
XX DE Sequence of 740 Arg-1649 Glu human Factor VIII:C.
XX
XX Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
XX KW haemophilia A.
XX
XX OS Homo sapiens.
XX
XX PN EP306968-A.
XX
XX PD 15-MAR-1989.
XX
XX PF 09-SEP-1988; 88EP-0114769.
XX
XX PR 10-SEP-1987; 87JP-0225147.
XX PR 08-APR-1988; 88JP-0085454.
XX
XX PA (KAGA ) CHERO SERO THERAPEUTIC RES INST.
XX PA (TEIJ ) TEIJIN LTD.
XX
XX Sugiyama T, Masuda K, Tajima Y, Yonemura H;
XX WPI; 1989-078467/11.
XX DR N-PSDB; AAN90654.
XX
XX Prodn. of recombinant human Factor-VIII-C -
XX PT using animal cells transformed with a vector contg. the gene for

```


PT Factor VIII:C and a promoter
XX
PS Disclosure: Fig 1; 32pp; English.
XX
CC Arg-740 of the carboxyl terminus of the H chain is
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of
CC L chain. A prefd. expression vector used to transform animal cell so
CC that they produce human Factor VIII:Cis plasmid Ad.RE.neo.
CC The expression vector has at least one promoter upstream of AAN90654.
CC The transformants can constantly and continuously produce human Factor
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so
CC produced is considered to corresp. to the smallest species of active and
CC intact Factor VIII:C molecules in the human blood plasma. It is useful
CC for treating haemophilia A patients.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1424 AA:

Query Match 93.4%; Score 127; DB 10; Length 1424;
Best Local Similarity 96.0%; Pred. No. le-09;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY I RPNYSRRLPKGVKHLKDFPILPGEI 25
|| |||||
Db 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

Search completed: October 17, 2003, 11:35:08
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:33:11 ; Search time 40 seconds
(without alignments)
60.105 Million cell updates/sec

Title: SEQID2-NAT3

Perfect score: 136

Sequence: 1 RPNYSRRLPKGVKHLKDFPILPGEI 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	93.4	2351	1 EZHU	coagulation factor
2	94	69.1	2319	2 A47004	coagulation factor
3	75	55.1	2133	2 T42763	coagulation factor
4	51	37.5	1607	2 T04583	TWV resistance pro
5	50	36.8	958	2 A82583	conserved hypotet
6	49	36.0	412	2 G71070	hypothetical prote
7	48.5	35.7	405	2 JC4591	alpha-1,3-fucosyl
8	48.5	35.7	405	2 B36340	alpha-1,3-fucosyl
9	48.5	35.7	433	2 A57596	alpha-1,3-fucosyl
10	48.5	35.7	1253	2 T45787	disease resistance
11	48	35.3	156	2 B83164	conserved hypotet
12	48	35.3	284	2 AC2515	hypothetical prote
13	48	35.3	600	2 B45642	DNA-directed DNA p
14	48	35.3	1690	2 T40847	probable rRNA biop
15	47	34.6	444	2 S54011	1-aminocyclopropan
16	46	33.8	308	2 S67857	hypothetical prote
17	46	33.8	348	2 T31681	hypothetical prote
18	46	33.8	467	2 T38353	serine hydroxymeth
19	46	33.8	480	2 A99164	hypothetical prote
20	46	33.8	519	2 T77572	oligopeptide trans
21	46	33.8	885	2 T38387	hypothetical prote
22	46	33.8	1001	2 C88779	protein T20D3.9 (i
23	46	33.8	1023	2 A59431	KIAA0013 protein [
24	46	33.8	1038	2 T25033	hypothetical prote
25	45.5	33.5	161	2 F75633	hypothetical prote
26	45.5	33.5	246	2 H82553	3-demethylubiquin
27	45.5	33.5	554	2 B90679	3-(3-hydroxyphenyl
28	45.5	33.5	554	2 P85528	3-(3-hydroxyphenyl
29	45.5	33.5	554	2 C64762	probable monooxyge

ALIGNMENTS

RESULT 1

EZHU

coagulation factor VIII precursor [validated] - human
N: Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant c
C: Species: Homo sapiens (man)
C: Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000
C: Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445;
R: Gitschier, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A: Title: Sequence of the exon-containing regions of the human factor VIII gene.
A: Reference number: I54318; MUID:93265012; PMID:1303178
A: Accession: I54318
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: DNA
A: Residues: 1-1921, 'S', 1923-2351 <RES>
A: Cross-references: GB:M88648; NID:gi82381; PIDN:AAA52420.1; PID:gi82383
R: Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; See
Nature 312, 330-337, 1984
A: Title: Expression of active human factor VIII from recombinant DNA clones.
A: Reference number: A00525; MUID:85061548; PMID:6438526
A: Accession: A00525
A: Molecule type: mRNA
A: Residues: 1-2351 <MOO>
A: Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R: Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.
S. D.N., Hewick, R.M.
Nature 312, 342-347, 1984
A: Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
A: Reference number: I58059; MUID:85061550; PMID:6438528
A: Accession: I58059
A: Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ
A: Molecule type: mRNA
A: Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>
A: Cross-references: GB:X01740; NID:gi82802; PIDN:AAA52484.1; PID:gi82803
R: Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.;
B.; Randolph, A.; Ordea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.
DNA 4, 333-349, 1985
A: Title: Characterization of the polypeptide composition of human factor VIII:C and
A: Reference number: A23584; MUID:86081164; PMID:3935400
A: Accession: A23584
A: Molecule type: mRNA
A: Residues: 1-2351 <TRU>
A: Cross-references: GB:M14113; NID:gi82817; PIDN:AAA52485.1; PID:gi82818
R: Eaton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A: Title: Proteolytic processing of human factor VIII. Correlation of specific cleava-
ity.
A: Reference number: A26174; MUID:86159740; PMID:3082357
A: Accession: A26174
A: Molecule type: protein
A: Residues: 20-36, 392-399, 'X', 401-402, 1668-1678, 1709-1722, 'D', 1723-1725, 1741-1755 <E
R: Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

30 45 33.1 61 2 AC2890
31 45 33.1 204 2 AB2365
32 45 33.1 438 2 S76347
33 45 33.1 444 2 T15310
34 45 33.1 1690 1 CGHULB
35 45 33.1 1691 1 CGH06B
36 45 33.1 1784 2 T10532
37 45 33.1 2054 2 T07584
38 44.5 32.7 331 2 F89771
39 44.5 32.7 485 2 G81308
40 44.5 32.7 788 2 A72330
41 44.5 32.7 906 2 JC5963
42 44 32.4 68 2 G69102
43 44 32.4 88 2 T17715
44 44 32.4 130 1 R3LW11
45 44 32.4 161 2 S33491

hypothetical prote
hypothetical prote
UDP-N-acetylglucos
hypothetical prote
collagen alpha 4(I
gag-pol polyprotei
hypothetical prote
lipoprotein (lipor
IMP dehydrogenase
phenylalanine-tRNA
stable tubule only
hypothetical prote
hypothetical prote
ribosomal protein
hypothetical prote

Biochemistry 31, 3315-3325, 1992

A:Title: Identification and functional importance of tyrosine sulfate residues within re

A:Reference number: A42348; MUID:92207952; PMID:1554716

A:Accession: A42348

A:Molecule type: protein

A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17

A:Experimental source: recombinant material from Chinese hamster ovary cells

A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi

R:Ray, P.J.; Smudzin, T.M.

J. Biol. Chem. 264, 14005-14010, 1989

A:Title: Intersubunit fluorescence energy transfer in human factor VIII.

A:Reference number: A43986; MUID:89340500; PMID:2503509

A:Accession: A43986

A:Molecule type: protein

A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>

R:Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.

J. Biol. Chem. 266, 740-746, 1991

A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for

A:Reference number: A56109; MUID:91093266; PMID:18987735

A:Contents: annotation; sulfation

R:Gitschier, J.; Wood, W.I.; Goralaka, T.M.; Wior, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,

Nature 312, 326-330, 1984

A:Title: Characterization of the human factor VIII gene.

A:Reference number: A56196; MUID:85063547; PMID:6438525

A:Contents: annotation; Introns

R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.

Protein Sci. 4, 740-746, 1995

A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o

A:Reference number: A56216; MUID:95338127; PMID:7613471

A:Contents: annotation; disulfide bonds

A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfinhydryls

R:Kjalke, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.

Eur. J. Biochem. 234, 773-779, 1995

A:Title: Amino acid residues 721-729 are required for full factor VIII activity.

A:Reference number: S63527; MUID:96163459; PMID:8575434

A:Accession: S63527

A:Molecule type: protein

A:Residues: 733-752;753-759 <KJA>

R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baerckman, M.; Almstedt, A.; Gray, E.; Sandberg

Eur. J. Biochem. 232, 19-27, 1995

A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction

A:Reference number: S66445; MUID:96048024; PMID:7556150

A:Accession: S66445

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1668-1685 <LIN>

C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr

C:Gene: GDB:F8C

A:Cross-references: GDB:119124; OMIM:306700

A:Map position: Xq28-Xq28

A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63

C:Function:

A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Ixa pro

A:Pathway: blood coagulation

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase

C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pia

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-2351/Product: coagulation factor VIII #status experimental <MAT>

F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>

F:20-356/Domain: A1 <DAL>

F:23-348/Domain: A2 <DA2>

F:392-759/Domain: A2 <DA2>

F:402-730/Domain: ferroxidase repeat homology <FO2>

F:760-1667/Domain: B <DBO>

F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <AC1>

F:1709-2038/Domain: A3 <DA3>

F:1716-2038/Domain: ferroxidase repeat homology <FO3>

F:2039-2191/Domain: C1 <DC1>

F:2039-2188/Domain: discoidin I amino-terminal homology <DNI>

F:2192-2351/Domain: C2 <DC2>

F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>

F:60,256,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198;267-348;547-573;649-730;1851-1877;1918-1922;2040-2188/Disulfide bonds: #s

F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted

F:365;737;738;742;1683;1699/Binding site: sulfate (Tyr) (covalent) #status experimen

F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimen

F:414;426/Binding site: sulfate (Tyr) (covalent) #status predicted

F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimen

F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental

F:1678-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experim

F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental

F:2193-2345/Disulfide bonds: #status predicted

Query Match 93.4%; Score 127; DB 1; Length 2351;

Best Local Similarity 96.0%; Pred. No. 2.6e-10; Gaps 0;

Matches 24; Conservative 0; Mismatches 1; Indels 0;

QY 1 RPNYSRRRLPGKGVKHLKDFPILPGEI 25

Db 503 RPLYSRRRLPGKGVKHLKDFPILPGEI 527

RESULT 2

A47004

coagulation factor VIII precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999

C:Accession: A47004

R:Elder, B.; Lakiich, D.; Gitschier, J.

Genomics 16, 374-379, 1993

A:Title: Sequence of the murine factor VIII cDNA.

A:Reference number: A47004; MUID:93300511; PMID:8314577

A:Accession: A47004

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2319 <ELD>

A:Cross-references: GB:I05573; MID:9162456; PIDN:AAA37385.1; PID:ql92457

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi

F:1-19/Domain: signal sequence #status predicted <SIG>

F:23-349/Domain: ferroxidase repeat homology <FO1>

F:402-730/Domain: ferroxidase repeat homology <FO2>

F:1686-2006/Domain: ferroxidase repeat homology <FO3>

F:2007-2156/Domain: discoidin I amino-terminal homology <DNI>

F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 69.1%; Score 94; DB 2; Length 2319;

Best Local Similarity 66.7%; Pred. No. 2.2e-05; Gaps 0;

Matches 16; Conservative 5; Mismatches 3; Indels 0;

QY 2 PNYSSRLPGKGVKHLKDFPILPGEI 25

Db 504 PLHARRLRPGIKHVRDLPIHPGEI 527

RESULT 3

T42763

coagulation factor VIII precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

C:Accession: T42763

R:Loillar, P.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z22269

A:Accession: T42763

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-2133 <LOI>

A:Cross-references: EMBL:049517; NID:gl511633; PID:gl511634; PIDN:AAB06705.1

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi

C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-2133/Product: coagulation factor VIII #status predicted <MAT>

F:23-349/Domain: ferroxidase repeat homology <FO1>

F:402-730/Domain: ferroxidase repeat homology <FO2>

F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Biochemistry 31, 3315-3325, 1992

A:Title: Identification and functional importance of tyrosine sulfate residues within re

A:Reference number: A42348; MUID:92207952; PMID:1554716

A:Accession: A42348

A:Molecule type: protein

A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17

A:Experimental source: recombinant material from Chinese hamster ovary cells

A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi

R:Ray, P.J.; Smudzian, T.M.

J. Biol. Chem. 264, 14005-14010, 1989

A:Title: Intersubunit fluorescence energy transfer in human factor VIII.

A:Reference number: A43986; MUID:89340500; PMID:2503509

A:Accession: A43986

A:Molecule type: protein

A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>

R:Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.

J. Biol. Chem. 266, 740-746, 1991

A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for

A:Reference number: A56109; MUID:91053266; PMID:18987735

A:Contents: annotation; sulfation

R:Gitschier, J.; Wood, W.I.; Goralaka, T.M.; Wior, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,

Nature 312, 326-330, 1984

A:Title: Characterization of the human factor VIII gene.

A:Reference number: A56196; MUID:85063547; PMID:6438525

A:Contents: annotation; Introns

R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.

Protein Sci. 4, 740-746, 1995

A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o

A:Reference number: A56216; MUID:95338127; PMID:7613471

A:Contents: annotation; disulfide bonds

A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfinhydryls

R:Kjalke, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.

Eur. J. Biochem. 234, 773-779, 1995

A:Title: Amino acid residues 721-729 are required for full factor VIII activity.

A:Reference number: S63527; MUID:96163459; PMID:8575434

A:Accession: S63527

A:Molecule type: protein

A:Residues: 733-752;753-759 <KJA>

R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baerckman, M.; Almstedt, A.; Gray, E.; Sandberg

Eur. J. Biochem. 232, 19-27, 1995

A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction

A:Reference number: S66445; MUID:96048024; PMID:7556150

A:Accession: S66445

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1668-1685 <LIN>

C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr

C:Gene: GDB:F8C

A:Cross-references: GDB:119124; OMIM:306700

A:Map position: Xq28-Xq28

A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63

C:Function:

A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Ixa pro

A:Pathway: blood coagulation

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase

C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pia

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-2351/Product: coagulation factor VIII #status experimental <MAT>

F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>

F:20-356/Domain: A1 <DAL>

F:23-348/Domain: A2 <DA2>

F:392-759/Domain: A2 <DA2>

F:402-730/Domain: ferroxidase repeat homology <FO1>

F:760-1667/Domain: B <DBO>

F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <AC1>

F:1709-2038/Domain: A3 <DA3>

F:1716-2038/Domain: ferroxidase repeat homology <FO3>

F:2039-2191/Domain: C1 <DC1>

F:2039-2188/Domain: discoidin I amino-terminal homology <DNI>

F:2192-2351/Domain: C2 <DC2>

F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>

F:60,256,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198;267-348;547-573;649-730;1851-1877;1918-1922;2040-2188/Disulfide bonds: #s

F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted

F:365;737;738;742;1683;1699/Binding site: sulfate (Tyr) (covalent) #status experimen

F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimen

F:414;426/Binding site: sulfate (Tyr) (covalent) #status predicted

F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimen

F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental

F:1678-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experim

F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental

F:2193-2345/Disulfide bonds: #status predicted

Query Match 93.4%; Score 127; DB 1; Length 2351;

Best Local Similarity 96.0%; Pred. No. 2.6e-10; Gaps 0;

Matches 24; Conservative 0; Mismatches 1; Indels 0;

QY 1 RPNYSRRRLPGKGVKHLKDFPILPGEI 25

DB 503 RPLYSRRRLPGKGVKHLKDFPILPGEI 527

RESULT 2

A47004

coagulation factor VIII precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999

C:Accession: A47004

R:Elder, B.; Lakiich, D.; Gitschier, J.

Genomics 16, 374-379, 1993

A:Title: Sequence of the murine factor VIII cDNA.

A:Reference number: A47004; MUID:93300511; PMID:8314577

A:Accession: A47004

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2319 <ELD>

A:Cross-references: GB:I05573; MID:9162456; PIDN:AAA37385.1; PID:ql92457

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi

F:1-19/Domain: signal sequence #status predicted <SIG>

F:23-349/Domain: ferroxidase repeat homology <FO1>

F:402-730/Domain: ferroxidase repeat homology <FO2>

F:1686-2006/Domain: ferroxidase repeat homology <FO3>

F:2007-2156/Domain: discoidin I amino-terminal homology <DNI>

F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 69.1%; Score 94; DB 2; Length 2319;

Best Local Similarity 66.7%; Pred. No. 2.2e-05; Gaps 0;

Matches 16; Conservative 5; Mismatches 3; Indels 0;

QY 2 PNYSSRLPGKGVKHLKDFPILPGEI 25

DB 504 PLHARRLRPGIKHVRDLPIHPGEI 527

RESULT 3

T42763

coagulation factor VIII precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

C:Accession: T42763

R:Loillar, P.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z22269

A:Accession: T42763

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-2133 <LOI>

A:Cross-references: EMBL:049517; NID:gl511633; PID:gl511634; PIDN:AAB06705.1

C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi

C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-2133/Product: coagulation factor VIII #status predicted <MAT>

F:23-349/Domain: ferroxidase repeat homology <FO1>

F:402-730/Domain: ferroxidase repeat homology <FO2>

F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Qy 1 RPNYSRRLPKGVKHLKDFILP 22

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:33:06 ; Search time 23 Seconds
(without alignments)

51.116 Million cell updates/sec

Title: SEQID2-NAT3

Perfect score: 136

Sequence: 1 RFRYSRLPKGVKHLKDPILPGEI 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	93.4	2351	1	FAB_HUMAN
2	94	69.1	2319	1	FAB_MOUSE
3	75	55.1	2133	1	FAB_PIG
4	54.5	40.1	239	1	UBIG_XANAC
5	54.5	40.1	239	1	UBIG_XANCP
6	52	38.2	598	1	COE4_MOUSE
7	48.5	35.7	405	1	FUT4_HUMAN
8	48.5	35.7	433	1	FUT4_MOUSE
9	48	35.3	495	1	DPO2_RAT
10	48	35.3	600	1	DPO2_MOUSE
11	46	33.8	467	1	GLYD_SCHPO
12	45.5	33.5	246	1	UBIG_XYLFA
13	45.5	33.5	554	1	MHPA_ECOLI
14	45	33.1	438	1	MURA_SYNY3
15	45	33.1	1032	1	IT44_XENLA
16	45	33.1	1690	1	CA44_HUMAN
17	45	33.1	1691	1	CA64_HUMAN
18	45	33.1	2054	1	YCF2_PINTH
19	44.5	32.7	788	1	SYF8_THEMEA
20	44	32.4	130	1	RR11_MARPO
21	44	32.4	300	1	ERA_MYCTU
22	44	32.4	356	1	CCDA_CAEEL
23	44	32.4	385	1	VATC_METTH
24	44	32.4	1143	1	CALI1_HUMAN
25	43.5	32.0	251	1	RPC8_YEAST
26	43.5	32.0	352	1	AROB_SULTO
27	43.5	32.0	649	1	VATI_CHLTR
28	43	31.6	138	1	RR11_PHAAN
29	43	31.6	146	1	YZ37_AQUAE
30	43	31.6	156	1	RS10_LUMRU
31	43	31.6	165	1	RS10_XENLA
32	43	31.6	166	1	RS10 ICTPU
33	43	31.6	227	1	FLPA_PYRHO

RESULT 1

ID	FAB_HUMAN	STANDARD	PRT	2351 AA
AC	P00451			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Coagulation factor VIII precursor (Procoagulant component)			
DE	(Antihemophilic factor) (AHF).			
GN	F8 OR F8C.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86081164; PubMed=3935400;			
RA	Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,			
RA	Hazog K., Kuo C.H., Maslarsz F.R., Merryweather J.P., Najarian R.,			
RA	Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,			
RA	Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,			
RA	Nordfang O., Ezban M.			
RT	*Characterization of the polypeptide composition of human factor			
RT	VIII: C and the nucleotide sequence and expression of the human kidney			
RT	CDNA.*;			
RL	DNA 4:333-349(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95061548; PubMed=6438526;			
RA	Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,			
RA	Keyl B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,			
RA	Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;			
RT	*Expression of active human factor VIII from recombinant DNA clones.*;			
RL	Nature 312:330-337(1984).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85061550; PubMed=6438528;			
RA	Toolle J.J., Knopf J.L., Wozney J.M., Suitzman L.A., Beecker J.L.,			
RA	Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,			
RA	Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,			
RA	Hewick R.M.;			
RT	*Molecular cloning of a cDNA encoding human antihemophilic factor.*;			
RL	Nature 312:342-347(1984).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93265012; PubMed=1303178;			
RA	Gitschier J., Wood W.I.;			
RT	*Sequence of the exon-containing regions of the human factor VIII			
RT	gene.*;			
RL	Hum. Mol. Genet. 1:199-200(1992).			
RN	[5]			
RP	SEQUENCE OF 2064-2070 FROM N.A.			
RA	de Water N.S., Williams R., Browett P.J.;			
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SULFATION OF TYR-1699.			

34	43	31.6	441	1	CYSP_THEAN	P25781 theileria a
35	43	31.6	529	1	TLM_SALPO	P07989 salmoneila
36	43	31.6	529	1	TLM_SALTY	P40813 salmoneila
37	43	31.6	579	1	ILVB_SPIPL	P27858 spirulina p
38	43	31.6	689	1	SYM_HALN1	Q9hsa4 halobacteri
39	43	31.6	891	1	SYV_PYRAB	Q9uy55 pyrococcus
40	43	31.6	1082	1	NC03_RAT	Q9epu2 rattus norv
41	43	31.6	1664	1	RPAL_YEAST	P10964 saccharomyc
42	43	31.6	3866	1	HRX_MOUSE	P59200 mus musculu
43	42.5	31.2	145	1	RS15_XIPMA	P70056 xiphophorus
44	42.5	31.2	188	1	CR12_HUMAN	P51864 homo sapien
45	42.5	31.2	203	1	RS4_CHLTE	P59129 chlorobium

ALIGNMENTS

RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of tyrosine 1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor.";
 RI J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP SULFATION
 RP MEDLINE=92207952; PubMed=1554716;
 RX Pittman D.D., Wang J.H., Kaufman R.J.;
 RA "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII.";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX Gilbert G.E., Baleja J.D.;
 RA "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy.";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A.";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A.";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT HEMA GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophiliac.";
 RL Science 232:1415-1416(1986).
 RN [13]
 RP VARIANT HEMA PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences.";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT HEMA GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides.";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT HEMA GLY-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Youssoufian H., Wong C., Kochan L., Olek K., Aronis S., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene.";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT HEMA CYS-1708.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";

RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A.";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT HEMA CYS-391.
 RX MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT HEMA LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT HEMA LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene.";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HEMA HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT HEMA CYS-1708.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1683 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A.";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANTS HEMA GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent.";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT HEMA CYS-391.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CNR haemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site.";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANTS HEMA PHE-1699 AND CYS-1708.
 RX MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA.";
 RL Genomics 6:65-71(1990).
 RN [25]
 RP VARIANTS HEMA CYS-1728 AND ASP-1941.
 RX MEDLINE=90169986; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";

FT	CARBOHYD	259	259	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	601	601	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	929	929	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	985	985	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1025	1025	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1111	1111	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1181	1181	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1208	1208	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1245	1245	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1265	1265	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1335	1335	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1408	1408	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1611	1611	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CARBOHYD	1919	1919	N-LINKED (GLCNAC.	-)	(POTENTIAL)
FT	CONFLICT	713	713	N -> M (IN REF. 2).		
FT	CONFLICT	734	734	I -> T (IN REF. 2).		
FT	CONFLICT	792	792	G -> Q (IN REF. 2).		
FT	CONFLICT	1133	1131	E -> F (IN REF. 2).		
FT	CONFLICT	1191	1191	I -> L (IN REF. 2).		
FT	CONFLICT	1209	1209	R -> P (IN REF. 2).		
FT	CONFLICT	1437	1437	C -> G (IN REF. 2).		
FT	CONFLICT	1456	1456	F -> R (IN REF. 2).		
FT	CONFLICT	1539	1539	F -> R (IN REF. 2).		
FT	CONFLICT	1546	1546	Q -> N (IN REF. 2).		
SQ	SEQUENCE	2133 AA:	239304 MW;	152BBA9997F570DA	CRC64;	
 Query Match 55.1%; Score 75; DB 1; Length 2133; Best Local Similarity 83.3%; Pred. No. 0 0052; Matches 15; Conservative 0; Mismatches 3; Indels 0;						
Qy	7	RLPKGVKHLKDPILPGE 24				
Dd	509	RLIKGWKLKDMPIIPGE 526				
 RESULT 4						
ID	UBIG_XANAC	STANDARD:	PRT;	239	AA.	
DT	QBPGK00;					
DC	28-FEB-2003 (Rel. 41, Created)					
DD	28-FEB-2003 (Rel. 41, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DE	3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-					
DE	dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHBB					
DE	methyltransferase).					
GN	UBig OR XAC2377.					
OS	Xanthomonas axonopodis (pv. citri).					
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;					
OC	Xanthomonadaceae; Xanthomonas.					
OX	NCBI_TaxID=92829;					
RP	[1]					
RN	SEQUENCE FROM N.A.					
RC	STRAIN=306 / ATCC 13902 / XV 101;					
RX	MEDLINE=22022145; PubMed=1204217;					
RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.					
RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.					
RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,					
RA	Camarote G., Cannavan F., Cardozo J., Chambergro F., Ciapina L.F.					
RA	Cacarelli R.M.B., Coutinho L.L., Cursivo-Santos J.R., El-Dorfi F.					
RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,					
RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,					
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.					
RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.					
RA	Martins E.C., Meidanus J., Menck C.F.M., Miyaki C.Y., Moon D.H.					
RA	Moreira L.M., Rossi M.T.M., Okura V.K., Oliveira M.C., Oliveira V.					
RA	Pereira H.A., Rossio A.D., Sena J.A.D., Silva C., de Souza R.F.,					
RA	Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza F.					
RA	Tindade dos Santos M., Truffi D., Tsai S.M., White F.F.,					
RA	Setubal J.C., Kitajima J.P.;					
RI	"Comparison of the genomes of two Xanthomonas pathogens with di					
RF	host specificities."					
RL	Nature 417:459-463(2002).					
QC	-1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-					

```

CC demethylubiquinone-9 - S-adenosyl-L-homocysteine + ubiquinone-9.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE011874; AAM37229.1; -.
CC HAMAP: MF_00472; 1.
CC InterPro: IPR001601; Methyltransf.
CC InterPro: IPR000051; SAM_bind.
CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
CC Complete proteome.
CC SEQUENCE 239 AA; 26060 MW; F56121516B27DFE6 CRC64;
CC -----
CC Query Match 40.1%; Score 54.5; DB 1; Length 239;
CC Best Local Similarity 57.1%; Pred. No. 0.58;
CC Matches 12; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
CC -----
CC QY 5 SRRLPKGVKHLKDFPILPGEI 25
CC :| |||| | || | | | |
CC Db 177 ARLLPKGTHHYKDF-IKPAEL 196
CC -----
CC RESULT 5
CC UBIG_XANCP STANDARD; PRT: 239 AA.
CC ID UBIG_XANCP STANDARD; PRT: 239 AA.
CC AC Q8P8H2;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3.4-
CC dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
CC methyltransferase).
CC DE UBIG OR XCC2269.
CC OS Xanthomonas campestris (pv. campestris).
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
CC Xanthomonadaceae; Xanthomonas.
CC OX NCBI_TaxID=340;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=2022145; PubMed=12024217;
CC RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
CC Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
CC Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
CC Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
CC Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Corri H.,
CC Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.L.T.,
CC Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
CC Katsuyama A.M., Kishi L.F., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
CC Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
CC Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moch D.H.,
CC Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
CC Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
CC Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.L.D.,
CC Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
CC Setubal J.C., Kitajima J.P.;
CC RT "Comparison of the genomes of two Xanthomonas pathogens with differing
CC host specificities."
CC RL Nature 417:459-463(2002)
CC CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 - S-adenosyl-L-homocysteine + ubiquinone-9.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE012334; AAM41548.1; -.
CC HAMAP: MF_00472; 1.
CC InterPro: IPR001601; Methyltransf.
CC InterPro: IPR000051; SAM_bind.
CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
CC Complete proteome.
CC SEQUENCE 239 AA; 26039 MW; 70BA381BAF14AC71 CRC64;
CC -----
CC Query Match 40.1%; Score 54.5; DB 1; Length 239;
CC Best Local Similarity 57.1%; Pred. No. 0.58;
CC Matches 12; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
CC -----
CC QY 5 SRRLPKGVKHLKDFPILPGEI 25
CC :| |||| | || | | | |
CC Db 177 ARLLPKGTHHYKDF-IKPSL 196
CC -----
CC RESULT 6
CC COE4_MOUSE STANDARD; PRT: 599 AA.
CC ID COE4_MOUSE STANDARD; PRT: 599 AA.
CC AC Q8K4J2; Q8K4J1; Q8K4J3; Q8K4J4; Q8K4J5;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Transcription factor COE4 (Early B-cell factor 4) (EBF-4) (Olf-1/EBF-
CC like 4) (OE-4) (O/E-4).
CC CN EBF4 OR COE4.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), AND TISSUE
CC SPECIFICITY.
CC RC STRAIN=C57BL/6;
CC RX MEDLINE=22136010; PubMed=12139918;
CC RA Wang S.S., Betz A.G., Reed R.R.;
CC RT "Cloning of a novel Olf-1/EBF-like gene, O/E-4, by degenerate oligo-
CC based direct selection."
CC RL Mol. Cell. Neurosci. 20:404-414(2002).
CC CC -!- FUNCTION: Seems to weakly activates transcription. Binds an Olf-1
CC consensus site in vitro.
CC CC -!- SUBUNIT: Forms either a homodimer or a heterodimer with a related
CC family member.
CC CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=3; Synonyms=4-23;
CC IsoId=Q8K4J2-1; Sequence=Displayed;
CC Name=1; Synonyms=4-11;
CC IsoId=Q8K4J2-2; Sequence=VSP_001125, VSP_001126;
CC Name=2; Synonyms=4-14;
CC IsoId=Q8K4J2-3; Sequence=VSP_001121, VSP_001122;
CC Name=4; Synonyms=4-132;
CC IsoId=Q8K4J2-4; Sequence=VSP_001123, VSP_001124;
CC Name=5; Synonyms=45;
CC IsoId=Q8K4J2-5; Sequence=VSP_001119, VSP_001120;
CC CC -!- TISSUE SPECIFICITY: Expressed in the neuronal and basal cell
CC layers of olfactory epithelium. Absent in the vomeronasal organ.
CC CC -!- SIMILARITY: BELONGS TO THE COE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

[2] SEQUENCE FROM N.A.
RP MEDLINE-91084663; PubMed-1702034;
RA Goelz S.E., Hession C., Goff D., Griffiths B., Tizard R., Newman B.,
RA Chi-Rosso G., Lobb R.;
RT "ELFT: a gene that directs the expression of an ELAM-1 ligand";
RL Cell 63:1349-1356(1990).
RN [3]
RP SEQUENCE OF 1-400 FROM N.A.
RX MEDLINE-92042084; PubMed-1718993;
RA Kumar R., Potvin B., Muller W.A., Stanley P.;
KT "Cloning of a human alpha(1,3)-fucosyltransferase gene that encodes
RT ELFT but does not confer ELAM-1 recognition on Chinese hamster ovary
cell transfectants.";
RI J. Biol. Chem. 266:21777-21783(1991).
RL -!- FUNCTION: May catalyze alpha-1,3 glycosidic linkages involved in
CC the expression of Lewis X/SSA-1 and VIM-2 antigens.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
CC form in trans cisternae of Golgi.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL: M65030; AAA92977.1; .
DR EMBL: M58596; AAB63172.1; .
DR EMBL: M58597; AAB63173.1; ALT_INIT.
DR EMBL: M65161; AAB20349.1; .
DR PIR: B36340; B36340.
DR Genew: HGNC:4015; FUT4.
DR MIM: 104230; .
DR GO: GO:0005624; C:membrane fraction; TAS.
DR GO: GO:0008417; F:fucosyltransferase activity; TAS.
DR GO: GO:0005975; P:carbohydrate metabolism; TAS.
DR InterPro: IPR001503; Glyco_trans_10.
DR Pfam: PF00852; Glyco_transf_1.
DR Transferase: Glycosyltransferase; Transmembrane; Glycoprotein;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSFEM 23 47 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 48 405 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 91 91 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 87 87 P -> R (IN REF. 2 AND 3).
FT CONFLICT 241 241 E -> D (IN REF. 3).
SQ SEQUENCE 405 AA; 45569 MW; DE72EIFDC390268D CRC64;

Query Match 35.7%; Score 48.5; DB 1; Length 405;
Best Local Similarity 50.0%; Pred. No. 8.6;
Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;
QY 1 RPNSRRRLPKGV-KHLKDFP 19
DB 319 RANRYEFVPRGAFIHVDFF 338

RESULT 8
FUT4_MOUSE
ID FUT4_MOUSE STANDARD; PRT: 433 AA.
AC Q11127;
DT 01-OCT-1996 (Rel. 34, Created)
JT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 4) (FUCT-IV);
DE cell adhesion.";

```
GN FUT4 OR ELFT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=96027607; PubMed=7559635;
RA Gersten K.M., Natsuka S., Trincherà M., Petryniak B., Kelly R.J.,
RA Hiraiwa N., Jenkins N.A., Gilbert D.J., Copeland N.G., Lowe J.R.;
RT "Molecular cloning, expression, chromosomal assignment, and tissue-
RT specific expression of a murine alpha-(1,3)-fucosyltransferase locus
RT corresponding to the human ELAM-1 ligand fucosyl transferase.";
RL J. Biol. Chem. 270:25047-25056(1995).
RN 2
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=97037075; PubMed=8882722;
RA Ozawa M., Muramatsu T.;
RT "Molecular cloning and expression of a mouse alpha-1,3
RT fucosyltransferase gene that shows homology with the human alpha-1,3
RT fucosyltransferase IV gene.";
RL J. Biochem. 119:302-308(1996).
CC -!- FUNCTION: May catalyze alpha-1,3 glycosidic linkages involved in:
CC the expression of Lewis X/SSEA-1 and VIM-2 antigens.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
CC form in trans cisternae of Golgi.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q11127-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q11127-2; Sequence=VSP_001778;
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN STOMACH AND COLON. IT
CC ALSO EXPRESSED IN THE LUNG, TESTIS, UTERUS, SMALL INTESTINE AND TO
CC A LESSEER EXTENT IN SPLEEN, TESTIS, AND OVARY. PRESENT IN TRACE AMOUNTS IN
CC BRAIN, THYMUS, HEART, SMOOTH MUSCLE, KIDNEY AND BONE MARROW. NOT
CC FOUND IN LIVER, SALIVARY GLAND AND PANCREAS.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U33457; AAC52269.1;
CC EMBL: D63380; BAA09697.1;
CC EMBL: D63379; BAA09696.1;
CC PIR: A57596; A57596.
CC MGD: MGI:95594; Fut4.
CC InterPro: IPR001503; Glyco_trans_10.
CC Pfam: PF00852; Glyco_transf_10;
CC Transferase; Glycosyltransferase; transmembrane; Glycoprotein;
CC Signal-anchor; Golgi stack; Alternative splicing.
CC DOMAIN 1 52 CYTOPLASMIC (POTENTIAL);
CC TRANSMEM 53 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL);
CC DOMAIN 75 433 LUMENAL, CATALYTIC (POTENTIAL);
CC CARBOHYD 117 117 N-LINKED (GLCNAC...) (POTENTIAL);
CC CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL);
CC VARSPLIC 1 33 Missing (in Isoform Short).
CC FT 252 Q -> P (IN REF. 2);
CC FT CONFLICT 257 R -> Q (IN REF. 2);
CC FT CONFLICT 260 V -> E (IN REF. 2);
CC FT CONFLICT 273 R -> Q (IN REF. 2);
CC SEQUENCE 433 AA; 49481 MW; 2401822F02B5D021 CRC64;
Query Match 35.7%; Score 48.5; DB 1; Length 433;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 RPNYSRRLPKGV-KHLKDFP 19
| | | | | | | | | |
DB 347 RANIERFVPRGAFIHVDQFP 366

RESULT 9
DPO2_RAT
ID DPO2_RAT STANDARD; PRT: 495 AA.
AC C89043;
DI 15-JUL-1999 (Rel. 38, Created)
DI 15-JUL-1999 (Rel. 38, Last sequence update)
DI 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA polymerase alpha 70 kDa subunit (DNA polymerase subunit B;
DE (DNA polymerase subunit II) (Fragment).
GN POLA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX Popanda O., Flohr C., Thielmann H.W.;
RT "A mutation in the gene of subunit II of DNA polymerase alpha from
RT Novikoff cells is concomitant with altered physico-chemical
RT properties of the enzyme.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE AT THE EARLY STAGE OF
CC CHROMOSOMAL DNA REPLICATION BY COUPLING THE POLYMERASE
CC ALPHA/PRIMASE COMPLEX TO THE CELLULAR REPLICATION MACHINERY (BY
CC SIMILARITY).
CC -!- SUBUNIT: DNA POLYMERASE ALPHA-PRIMASE IS A FOUR SUBUNIT ENZYME
CC (SUBUNITS A, B, C AND D), WHICH IS ASSEMBLED THROUGHOUT THE CELL
CC CYCLE. THE LARGEST SUBUNIT (SUBUNIT A) HAS DNA POLYMERASE
CC ACTIVITY, THE TWO SMALLEST SUBUNITS (SUBUNITS C AND D) HAVE DNA
CC PRIMASE ACTIVITY. SUBUNIT B BINDS TO SUBUNIT A.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: PHOSPHORYLATED IN A CELL CYCLE-DEPENDENT MANNER, IN G2/M
CC PHASE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE ALPHA SUBUNIT B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ011606; CAA09721.1;
CC Pfam: PF04058; DNA_pol_alpha_B;
CC DNA replication; Nuclear protein; Phosphorylation.
CC NON_TER 1 1
CC DOMAIN 10 52 PRO/SER/THR-RICH (HYDROPHILIC).
CC SEQUENCE 495 AA; 54952 MW; 1816F035737C4A0B CRC64;
Query Match 35.3%; Score 48; DB 1; Length 495;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 7 RLPKGVKHLKDFPILPGEI 25
:::| | | | | | | | |
DB 178 QIPVDVSELKDYSLPFGV 196

RESULT 10
DPO2_MOUSE
ID DPO2_MOUSE STANDARD; PRT: 600 AA.
AC P33611;
DI 01-FEB-1994 (Rel. 28, Created)
```

DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA polymerase alpha 70 kDa subunit (DNA polymerase subunit B).
 GN POLA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 84-102; 269-265 AND 354-403.
 RX MEDLINE=93216788; PubMed=8463324;
 RA Miyazawa H., Izumi M., Tada S., Takada R., Masutani M., Uli M.,
 RA Hanacka F.;
 RT "Molecular cloning of the cDNAs for the four subunits of mouse DNA
 RT polymerase alpha-prime complex and their gene expression during
 RT cell proliferation and the cell cycle.";
 RL J. Biol. Chem. 268:8111-8122(1993).
 CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE AT THE EARLY STAGE OF
 CC CHROMOSOMAL DNA REPLICATION BY COUPLING THE POLYMERASE
 CC ALPHA-PRIMASE COMPLEX TO THE CELLULAR REPLICATION MACHINERY (BY
 CC SIMILARITY).
 CC -!- SUBUNIT: DNA POLYMERASE ALPHA-PRIMASE IS A FOUR SUBUNIT ENZYME
 CC (SUBUNITS A, B, C AND D), WHICH IS ASSEMBLED THROUGHOUT THE CELL
 CC CYCLE. THE LARGEST SUBUNIT (SUBUNIT A) HAS DNA POLYMERASE
 CC ACTIVITY. THE TWO SMALLEST SUBUNITS (SUBUNITS C AND D) HAVE DNA
 CC PRIMASE ACTIVITY. SUBUNIT B BINDS TO SUBUNIT A.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- PTM: PHOSPHORYLATED IN A CELL CYCLE-DEPENDENT MANNER, IN G2/M
 CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE ALPHA SUBUNIT 5 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D13546; BAA02746.1; -
 DR PIR: B46642; B46642
 DR MGD: MGI:99690; Pola2.
 DR Pfam: PF04058; DNA_pol_alpha_B; 1.
 KW DNA replication; Nuclear protein; Phosphorylation.
 FT DOMAIN 101 107 POLY-GLU.
 FT DOMAIN 115 157 PRO/SER/THR-RICH (HYDROPHILIC).
 SQ SEQUENCE 600 AA; 66267 MW; 79F94BE6EF33FEB3 CRC64;
 Query Match 35.3%; Score 48; DB 1; Length: 600;
 Best Local Similarity 26.1%; Pred. No. 16;
 Matches 6; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
 QY 3 NYSRRLPKGVKHLKDFPILPGEI 25
 Db 279 SYGAIPVDLSLKEYSLFPGOV 301
 RESULT 11
 ID GLYD_SCHPO STANDARD; PRI: 467 AA.
 AC O13972;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable serine hydroxymethyltransferase, cytosolic (EC 2.1.2.1);
 DE (serine methylase) (Glycine hydroxymethyltransferase) (SHMT).
 GN SPAC24C9.12C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras K., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrsach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Mambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RG Nature 415:871-880(2002).
 CC -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -!- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate - glycine +
 CC H(2)O -> tetrahydrofolate + L-serine.
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z58601; CAB11269.1; -
 DR PIR: T38353; T38353.
 DR HSSP: P07511; ICDU.
 DR GeneDB_SPombe; SPAC24C9.12C; -
 DR InterPro: IPR001085; Gly_HyMetTransf.
 DR Pfam: PF00464; SHMT; 1.
 DR PROSITE: PS00096; SHMT; 1.
 KW Transferase; Pyridoxal phosphate; One-carbon metabolism.
 FT BINDING 243 243 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT DOMAIN 238 241 POLY-THR.
 SQ SEQUENCE 467 AA; 51861 MW; 0C21D7EF010C3725 CRC64;
 Query Match 33.8%; Score 46; DB 1; Length 467;
 Best Local Similarity 45.5%; Pred. No. 24;
 Matches 10; Conservative 1; Mismatches 11; Indels 0; Gaps 0;
 QY 3 NYSRRLPKGVKHLKDFPILPGE 24
 Db 419 NLQKELPKDANKLKDFKALGE 440
 RESULT 12
 CBIG_XYLFA
 ID UBIG_XYLFA STANDARD; PRI: 246 AA.
 AC Q9PAM5;
 DT 28-FEB-2003 (Rel. 41, Created)

```
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHB
DE UBIG OR XF2471)
GN UBIG OR XF2471
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
ON NCBI_TaxID=2371;
RX SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Canarro A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa P.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraja J.S., Franca S.C., Franco M.C., Frome M., Furlan L.R.,
RA Garner J.M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hebeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.F., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.N., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A., Jr., Nobrega F.G., Nunes L.C., Oliveira M.A.,
RA de Oliveira M.C. de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshukko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 -> S-adenosyl-L-homocysteine + ubiquinone-9.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC004055; AAF85269.1; -
CC PIR: H82553; H82553.
CC HAMAP: MF_00472; - 1.
CC InterPro: IPR001601; Methyltransf.
CC InterPro: IPR000051; SAM bind.
CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
CC Complete proteome.
CC SEQUENCE 246 AA: 27121 MW: 5212107D3633D3F3 CRC64:
Query Match 33.5%: Score 45.5; DB 1; Length 246;
Best Local Similarity 55.0%: Pred. No. 14;
Matches 11; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
QY 6 RRLPKGVKHLKDFPILPGEI 25
| | | | |
Db 178 RLLPKGTHQYKDF-IRPAEL 196
```

```
RESULT 13
MHPA_ECOLI STANDARD; PRT: 554 AA.
ID AC P77397; P71203; P77047;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-(3-hydroxy-phenyl)propionate hydroxylase (EC 1.14.13.-).
GN MHPA OR B0347.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
ON NCBI_TaxID=562;
RX SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Kawamuka M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / CS520;
RA Ferrandez A., Garcia J.L., Diaz E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: FAD (BY SIMILARITY).
CC -!- PATHWAY: 3-hydroxyphenylpropionate degradation.
CC -!- SIMILARITY: BELONGS TO THE PHEA/TEFR FAMILY OF FAD MONOOXYGENASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D86239; BAAL3052.1; -
CC EMBL: Y09555; CAA70747.1; -
CC EMBL: AE000142; AAC73450.1; -
CC EMBL: U73857; AAB18071.1; -
CC PIR: C64762; C64762.
CC EcoGene: EG20273; mhpA.
CC InterPro: IPR000733; Flav_monooxygenase.
CC InterPro: IPR002938; Moxy_FAD_binding.
CC InterPro: IPR000205; NAD_binding.
CC InterPro: IPR003042; Rng_mnoxygenase.
CC Pfam: PF01494; FAD_binding_3; 1.
CC Pfam: PF01360; Monooxygenase; 1.
CC PRINTS: PR00420; RINGMONOXGNASE.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Flavoprotein; FAD;
CC Complete proteome.
CC NP_BIND 17 46 FAD (POTENTIAL).
CC NP_BIND 285 295 FAD (POTENTIAL).
CC CONFLICT 272 292 Q -> H (IN REF. 1).
CC CONFLICT 360 360 L -> P (IN REF. 1).
CC CONFLICT 502 502 W -> G (IN REF. 1).
CC SEQUENCE 554 AA: 62185 MW: 1D56CB799EF9A8E CRC64:
```

Query Match 33.5%; Score 45.5; DB 1; Length 554;
 Best Local Similarity 41.7%; Pred. No. 35;
 Matches 10; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 1 RPNYSRLPKGVKHLKDPILPGE 24
 || | | | : | : |||
 DB 220 RPYVSAALPHAVRRF-EFVMFGE 242

RESULT 14
 MURA_SYNV3 STANDARD; PRT: 438 AA.
 AC Q55573;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
 DE (Enolpyruvate transferase) (UDP-N-acetylglucosamine enolpyruvyl transferase) (EPT).
 GN MURA OR MURZ OR SLR0017.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N., Sugiura M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome.";
 RL DNA Res. 2:153-165(1995).
 CC -1- FUNCTION: CELL WALL FORMATION. ADOS ENOLPYRUVYL TO UDP-N-ACETYLGLUCOSAMINE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + UDP-N-acetyl-D-glucosamine = phosphate + UDP-N-acetyl-3-(1-carboxyvinyl)-D-glucosamine.
 CC -1- PATHWAY: Peptidoglycan biosynthesis; first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: Belongs to the EPSP synthase family. Mura subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D64000; BAA10199.1; -;
 CC PIR: S76347; S76347.
 CC DR HSSP: P33038; IDLG.
 CC DR HAMAP: ME_00111; -; 1.
 CC DR InterPro: IPR005750; AcGlu_Trans_Mura.
 CC DR InterPro: IPR001986; EPSP_Synthase.
 CC DR Pfam: PF00275; EPSP_Synthase; 1.
 CC DR ProDom: PD001867; EPSP_Synthase; 1.
 CC DR TIGRFAMs: TIGR01072; mura; 1.
 CC KW Peptidoglycan synthesis; Cell wall; Cell division; Transferase;
 KW Complete proteome.
 FT ACT_SITE 129 129 BINDS PER (BY SIMILARITY).
 SQ SEQUENCE 438 AA; 46604 MW; C310AB287F2F0F55 CRC64;

Query Match 33.1%; Score 45; DB 1; Length 438;
 Best Local Similarity 47.8%; Pred. No. 32;
 Matches 11; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 5 SRRLPKGVK--HLKDPILPGEI 25
 || : || : | ||| : |
 DB 228 SRILINGVEKLNSTDFPIIPDRI 250

RESULT 15
 ITA4_XENLA STANDARD; PRT: 1032 AA.
 AC Q91687; Q06273;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-4 precursor (Integrin alpha-IV) (VLA-4).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96379747; PubMed=8787760;
 RA Ramos J.W., Whittaker C.A., Desimone D.W.;
 RT "Integrin-dependent adhesive activity is spatially controlled by inductive signals at gastrulation.";
 RL Development 122:2873-2883(1996).
 RN [2]
 RP SEQUENCE OF 308-379 FROM N.A.
 RX MEDLINE=94008528; PubMed=8404528;
 RA Whittaker C.A., Desimone D.W.;
 RT "Integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryos.";
 RL Development 117:1239-1249(1993).
 CC -1- FUNCTION: FIBRONECTIN AND V-CAM ADHESION RECEPTOR.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U54497; AAA58673.1; -;
 CC EMBL: L10188; AAA16248.1; -;
 CC PIR: I51526; I51526.
 CC DR HSSP: P11215; LA8X.
 CC DR InterPro: IPR000413; Integrin_alpha.
 CC DR Pfam: PF01839; FG-GAP; 3.
 CC DR Pfam: PF00357; Integrin_A; 1.
 CC DR PRINTS: PRO1185; INTEGRINA.
 CC DR SMART: SM00191; Int_alpha; 5.
 CC DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
 CC DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat.
 FT SIGNAL 1 34 POTENTIAL.
 FT CHAIN 35 1032 INTEGRIN ALPHA-4.
 FT DOMAIN 35 974 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 975 998 POTENTIAL.
 FT DOMAIN 999 1032 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 50 112 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT REPEAT 197 247 FG-GAP 3.
 FT REPEAT 248 301 FG-GAP 4.
 FT REPEAT 303 361 FG-GAP 5.
 FT REPEAT 364 423 FG-GAP 6.
 FT REPEAT 426 474 FG-GAP 7.
 FT CA_BIND 314 322 POTENTIAL.
 FT CA_BIND 376 384 POTENTIAL.
 FT CA_BIND 438 446 POTENTIAL.
 FT SITE 590 591 CLEAVAGE.
 FT SITE 1001 1005 GFFKR MOTIF.
 FT DISULFID 91 101 BY SIMILARITY.
 FT DISULFID 144 165 BY SIMILARITY.
 FT DISULFID 183 198 BY SIMILARITY.

FT	DISULFID	485	494	BY SIMILARITY.
FT	DISULFID	500	556	BY SIMILARITY.
FT	DISULFID	622	627	BY SIMILARITY.
FT	DISULFID	698	712	BY SIMILARITY.
FT	DISULFID	853	889	BY SIMILARITY.
FT	DISULFID	896	901	BY SIMILARITY.
FT	CARBOHYD	81	81	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	229	229	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	479	479	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	496	496	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	517	517	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	537	537	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	626	626	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	660	660	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	746	746	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1032 AA;	115215 MW;	6486797D93AAE63E CRC64;

Query Match 33.1%; Score 45; Db 1; Length 1032;
Best Local Similarity 52.6%; Pred. No. 83;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY	7	RLPKGVKHLKDPPILPGEI	25
		: :	
Db	678	RLPKGLYFVKVFDLLEKEI	696

Search completed: October 17, 2003, 11:33:39
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:33:11 ; Search time 95 Seconds
(without alignments)
67.909 Million cell updates/sec

Title: SEQID2-NAT3
Perfect score: 136
Sequence: 1 PNYSRRLPGVKVHLKDFILPGEI 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organellae:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	74.3	2343	6 O62730	O62730 canis fami
2	101	74.3	2343	6 O18806	O18806 canis fami
3	53.5	39.3	611	16 O8EYK5	O8EYK5 leptospira
4	51.5	37.9	911	16 O8DAC5	O8DAC5 vibrio vuln
5	51	37.5	166	4 O96AX6	O96AX6 homo sapien
6	51	37.5	907	16 O8DLH3	O8DLH3 synectococc
7	51	37.5	1607	10 O65506	O65506 arabidopsis
8	50.5	37.1	589	2 O9S158	O9S158 comamonas t
9	50.5	37.1	935	5 O9VE79	O9VE79 drosophila
10	50	36.8	388	10 O8SAW1	O8SAW1 oryza sativ
11	50	36.8	579	4 O96AB1	O96AB1 homo sapien
12	50	36.8	706	10 O8SSJ1	O8SSJ1 oryza sativ
13	50	36.8	958	16 O9PBA7	O9PBA7 xyella fas
14	50	36.8	2618	5 O9VPB8	O9VPB8 drosophila
15	49	36.0	412	17 O5B994	O5B994 pyrococcus
16	49	36.0	986	10 O8GRU6	O8GRU6 lotus japon

17	48.5	35.7	386	11	O920V9	mus musculu
18	48.5	35.7	390	11	O91V20	mus musculu
19	48.5	35.7	390	11	O920V7	mus musculu
20	48.5	35.7	390	11	O920V6	mus musculu
21	48.5	35.7	390	11	O920W1	mus musculu
22	48.5	35.7	390	11	O920W0	mus musculu
23	48.5	35.7	390	11	O91VB5	mus musculu
24	48.5	35.7	390	11	O920V8	mus musculu
25	48.5	35.7	711	10	O8RYK8	oryza sativ
26	48.5	35.7	1253	10	O9SC23	arabidopsis
27	48	35.3	55	7	O9MWB4	alcolapia a
28	48	35.3	156	16	O9HXP9	pseudomonas
29	48	35.3	264	16	O8YK16	anabaena sp
30	48	35.3	360	11	O922M1	mus musculu
31	48	35.3	454	11	O8C111	mus musculu
32	48	35.3	566	11	O8C2T6	mus musculu
33	48	35.3	600	11	O8VDR3	mus musculu
34	48	35.3	600	11	O9QYV6	rattus norv
35	48	35.3	1161	10	O9LR20	arabidopsis
36	48	35.3	1690	3	O74835	schizosacch
37	47.5	34.9	350	2	O93F54	pseudomonas
38	47.5	34.9	352	2	O93F60	pseudomonas
39	47.5	34.9	352	2	O93F48	pseudomonas
40	47.5	34.9	401	11	O93JK2	cricetulus
41	47.5	34.9	433	11	O99N88	rattus norv
42	47	34.6	187	16	O82NP8	salmonella
43	47	34.6	345	16	O8NLX9	corynebacte
44	47	34.6	444	10	O43756	doritaenops
45	47	34.6	537	16	O8E9K6	shewanella

ALIGNMENTS

RESULT 1
O62730 PRELIMINARY: PRT: 2343 AA.
ID O62730
AC O62730;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, and Spleen;
RA Gordy P.W., Bowen R.A.;
RT "Characterization of the canine factor VIII cDNA."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL: AF049489; AAC05384.1;
DR HSSP: P00451; ICTG
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR000421; FA58_C.
DR Pfam: PF00394; Cu-oxidase; 3.
DR Pfam: PF00754; F5_F8_type_C; 2.
DR SMART: SM00231; FA58C_2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD59 CRC64;

Query Match 74.3% Score 101; DB 6; Length 2343;
Best Local Similarity 79.2%; Pred. No. 3.3e-06;
Matches 19; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYSRRLPGVKVHLKDFILPGEI 25

DB 498 PLHTGLPGVKVHLKDFILPGEI 521

Qy	6 RRLP--KGV-KHLKDFPILPGEI 25
	:
Db	450 KRYPLIKGIFKNVSGFPILPGEV 472

```

DR Pfam; PF00560; LRR; 3.
DR Pfam; PF00931; NB-ARC; 2.
DR Pfam; PF01502; TIR; 2.
DR PRINIS; PR00364; DISEASERS1ST.
DR SMART; SM00382; AAA; 2.
DR SMART; SM00255; TIR; 2.
DR PROSITE; PS0104; TIR; 1.
KW ATP-binding.
SQ SEQUENCE 1607 AA; 182785 MW; 4F8F572EC72074F3 CRC64;

Query Match 37.5%; Score 51; DB 10; Length 1607;
Best Local Similarity 66.7%; Pred.No. 97;
Matches 10; Conservative 2; Mismatches 3; Indels 0;

Qy 3 NYSRRRLPGKGVKHLKD 17
   :|| |||||:| |
Db 573 HYSRLRPLKGLKLPD 987

RESULT 8
IC Q9S158 PRELIMINARY; PRT; 589 AA.
AC Q9S138:
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE 3-(3-hydroxyphenyl)propiolate hydroxylase.
GN MHPA.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
SC STRAIN=TA441;
XX MEDLINE=20005595; PubMed=10537203;
RA Arai H., Yamamoto T., Ohishi T., Shimizu T., Nakata T., Kudo T.
RT "Genetic organization and characterization of the 3-(3-
RT hydroxyphenyl)propiolate acid degradation pathway of Comamonas
RL testosteroni TA441."
RL Microbiology 145:2813-2820(1999).
DR EMBL; AB024335; BAA82878.1; -.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR000733; Flav_monooxygense.
DR InterPro; IPR002114; Hpr_SerP_site.
DR InterPro; IPR002938; MoxY_FAD_binding.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR001042; Hmg_mnnoxygenase.
DR InterPro; IPR000594; ThiF_domain.
DR Pfam; PF01494; FAD_binding_3; 1.
DR Pfam; PF01360; Monooxygenase; 1.
DR PRINTS; PR00369; FADPNR.
DR PRINTS; PR00420; RNMNOXGNASE.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
SQ SEQUENCE 589 AA; 65297 MW; 7CACE905825CB744 CRC64;

Query Match 37.1%; Score 50.5; DB 2; Length 589;
Best Local Similarity 41.7%; Pred.No. 38;
Matches 10; Conservative 5; Mismatches 8; Indels 1;

Qy 1 RPNYSRRRLPGKGVKHLKDFPILPGE 24
   || | || ||:| :|||
Db 218 RPYVSAALPHGIRRP-EFVWMPGE 240

RESULT 9
IC Q9VE79 PRELIMINARY; PRT; 935 AA.
AC Q9VE79:
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE CG14309 protein.

```

GN	CGI4309.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BERKELEY;	
RC	MEDLINE=20196006; PubMed=10731132;	
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Ananastos P.G., Scherer S.E., Li P.W., Hoskins R.A., Gaille R.F.,	
RA	George R.A., Lewis S.F., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Stratton R.C., Rogers J., Blake J.R., Champetier R.G., Champsin D.,	
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	
RA	Beeason K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,	
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,	
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov S.C., Dunc P.,	
RA	Durban K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,	
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	
RA	Hoskins N., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA	Mount K.G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	
RA	Mourou G., Murphy B., Murphy L., Wenz D.M., Nelson D.L.,	
RA	Neison D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,	
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao C., Zheng L.,	
RA	Zheng X.H., Zhong F.N., Zhong Y., Zhou X., Zhu S., Zhu X., Smith H.C.,	
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.	
RT	"The genome sequence of Drosophila melanogaster."	
DL	Science 287:2185-2195(2000).	
DR	EMBL: A5003721; AAF5548.1; --	
DR	FlyBase: FBgn0038611; AAF5548.1; --	
SQ	SEQUENCE 935 AAT 106833 MW 10E46B721CE5D83F CRC64;	

```
Query Match      37.1% Score 50.5; DB 5; Length 935;
Best Local Similarity 45.8%; Pred. No. 64;
Matches 11; Conservative 4; Mismatches 8; Indexes 1; Gaps 1;
```

RESULT	10
Q8SAW1	
ID	PRELIMINARY; PRT; 388 AA.
AC	Q8SAW1;
DT	01-JUN-2002 (TREMBLrel_21, Created)
DT	01-JUN-2002 (TREMBLrel_21, Last sequence update)
DT	01-MAR-2003 (TREMBLrel_23, Last annotation update)
DE	Putative wall-associated protein kinase.
GN	OSJNERA0051J07.1
OS	Oryza sativa (rice).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC	Eurhartoidea; Gryzeae; Oryza.
EN	NCBI_TaxID=4530;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Wing R.A., Yu Y.D., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA	Saski C.A., Henry D., Oates R., Simmons J.;
RT	"Rice Genomic Sequence."
FT	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AC098566; AAL77114.1; -
DR	Gramene: Q8SAML; -
DR	InterPro: IPR000152; Asx_hydroxyl.
DR	InterPro: IPR001881; EGF_CA.
DR	SMART: SM00179; EGF_CA; 1.
DR	PROSITE: PS00010; ASX_HYDROXYL; 1.
DR	PROSITE: PS01187; EGF_CA; 1.
KW	EGF-like domain; Kinase.
QY	SEQUENCE 388 AA: 42179 MW; 0F4F47E6F087FF91 CRC64;
Query Match 36.8%; Score 50; DB 10; Length 388;	
Best Local Similarity 47.6%; Pred. No. 29;	
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps	
Oy	3 NYSRRLPKGVKHLKDFPLPG 23 I:: :: :: :: :: :: :: ::
Dd	240 NFSKKYPKGVLVIDFAIRDG 260 I:: :: :: :: :: :: ::
RESULT 11	
Q96A81	
ID	Q96A81 PRELMINARY; PRT; 579 AA.
AC	Q36A81;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	Hypothetical protein FLJ31952.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA	Furuwa T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA	Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irif R.,
RA	Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA	Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA	Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA	Kanehori K., Takahashi-Fuji A., Oshima A., Sugiyama A., Kawakami B.
RA	Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isozai T.;
RT	"NEDO human cDNA sequencing project."
XL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AK056514; BAB71201.1; -
KW	Hypothetical protein.
QY	SEQUENCE 579 AA: 65949 MW; 195AEFB2F4235C84 CRC64;

```

Query Match      36.8% ; Score 50; DB 4; Length 579;
Best Local Similarity 40.0% ; Pred. No. 45;
Matches 12; Conservative 5; Mismatches 5; Indels 8; Gaps 2;

Qy 1 RPNYSRRLPKGVKHLKD-----FPILEGEEI 25
   |||:::|
Db 61 RPNYSK--KGLEHKADIQOHLFPVPPGHL 87

```

[illegible]

```

OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA Kuit K., Nascimento L., Zutaevern T., Balijs V., King M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RI "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RI OS7NB80011A24, from chromosome 10, complete sequence.";
RI Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC113336; AA001181.1;
DR EMBL; GR085531;
DR Gramene; Q85531;
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR000719; Pro_kinase.
DR Pfam; PF000069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00179; EGF_CA; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding, EGF-like domain; Kinase; Transrase.
SQ SEQUENCE 706 AA; 77555 MW; 91EGF07E9B90BFEB CRC64;

Query Match 36.8%; Score 50; DB 10; Length 706;
Best Local Similarity 47.6%; Pred. No. 56;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NYSRRLPKGVKHLKDFPILPG 23
Db 240 NFSKKYKGVPLVDFEIRDG 260

RESULT 13
Q9PB7 3 NYSRRLPKGVKHLKDFPILPG 23
ID Q9PB7 PRELIMINARY; PRT; 558 AA.
AC Q9PB7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Xf2237.
GN Xf2237.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=945C;
RA MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu P.A., Accencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Cartaro D.M., Carrer H.,
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docuna C., El-Dorzy H.,
RA Facincani A.P., Ferreira A.J.S., Ferrelira V.C.A., Fertiz J.A.,
RA Freiga J.S., Franca S.C., Franco M.C., Froume M., Furian L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lenos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madira H.M.F., Marino C.B.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

```

```

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RL "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000);
DR EMBL; AE004036; AAF85036.1;
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00593; TonB_dep_Rec; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 958 AA; 103030 MW; 4EF720E0D333EDB CRC64;

Query Match 36.8%; Score 50; DB 16; Length 958;
Best Local Similarity 47.1%; Pred. No. 78;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 RRLPKGVKHLKDFPILP 22
Db 124 RRLPKGVKHLKDFPILP 140

RESULT 14
Q5VP88 6 RRLPKGVKHLKDFPILP 22
ID Q5VP88 PRELIMINARY; PRT; 2618 AA.
AC Q5VP88; Q9N6F9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG9936 protein (JLKJLKJ) (TRAP240) (PAP/DTRAP240) (Transcriptional
DE coactivator BLIND spot).
GN PAP OR BLI OR CG9936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Hallett R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Jostli D., Kalush K.A., Howland I.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kaulston F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mourtoulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Munkit S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-E., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Ceiniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon K.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fartan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jaitani M., Kruse D., Li P., Mattel B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puti V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C.; Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.E.,
RA Tupy J.L., Bergman C., Berken B., Carlson J.W., Celisnik S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harts N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RA Boube M., Faucher C., Joulia L., Cribbs D.L., Bourbon H.M.;
RT "The *Drosophila* Mediator proteins DTRAP240 and DTRAP80 are
RT transcriptional cofactors of the proboscipedia and Sex combs reduced
RT homeotic proteins.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RA Nairz K., Hafen E.;
RT "Isolation of flytrap (pap), the *Drosophila* TRAP240 homologue.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RN SEQUENCE FROM N.A.
RA Boube M., Faucher C., Joulia L., Cribbs D.L., Bourbon H.M.;
RT "The *Drosophila* Mediator proteins DTRAP240 and DTRAP80 are
RT transcriptional cofactors of the proboscipedia and Sex combs reduced
RT homeotic proteins.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [9]
RN SEQUENCE FROM N.A.
RA Treisman J.E.;
RT "Drosophila homologs of the transcriptional coactivation complex
RT subunits TRAP240 and TRAP230 are required for identical processes in
RT eye-antennal disc development.";
RL Development 0:0-0(2001).
RA EMBL; AE003593; AAF51667.2; -

DR EMBL: AF227215; AAF43172.1; -;
DR EMBL: AF226855; AAF36691.1; -;
DR EMBL: AF227214; AAF43021.1; -;
DR EMBL: AF324425; AAG48327.1; -;
SQ FLYBASE: FBgn0024200; pap;
SQ SEQUENCE 2618 AA; 280021 MW; 735A9A502076844E CRC64;
Query Match 36.8%; Score 50; DB 5; Length 2618;
Best Local Similarity 40.9%; Pred. No. 2.4e-02;
Matches 9; Conservative 6; Mismatches 5; Indels 2; Gaps 1;
Oy 1 RPNYSRRLPKGVKHLKDFILP 22
Db 836 RPRYAKNLVEGQNHVK--PVMP 855
RESULT 15
ID O58994 PRELIMINARY; PRT; 412 AA.
AC O58994
DT 01-AUG-1998 (TRIMBLrel. 07, Created)
DT 01-AUG-1998 (TRIMBLrel. 07, Last sequence update)
DT 01-OCT-2002 (TRIMBLrel. 22, Last annotation update)
DE Hypothetical protein PH1259.
GN PH1259.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
EX MEDLINE-98344137; PubMed-9679194;
RA Kawarabayashi Y.; Sawada M.; Horikawa H.; Haikawa Y.; Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000005; BAA30361.1; -;
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR001566; Tima.
DR PROSITE: PS01231; TRMA_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 412 AA; 47094 MW; FF92CDD11EF5B331 CRC64;
Query Match 36.0%; Score 49; DB 17; Length 412;
Best Local Similarity 58.8%; Pred. No. 44;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Oy 2 PNYSRRLPKGVKHLKDF 18
Db 140 PIFSERTPKYLKALKDF 156
Search completed: October 17, 2003, 11:37:41
Job time : 99 secs

1	127	93.4	1438	4	US-09-209-916-1	Sequence 1, Appli
2	127	93.4	1471	1	US-08-683-835B-3	Sequence 3, Appli
3	127	93.4	1561	2	US-08-888-083-2	Sequence 2, Appli
4	127	93.4	1661	2	US-08-558-107-2	Sequence 2, Appli
5	127	93.4	1661	3	US-09-243-539-2	Sequence 2, Appli
6	127	93.4	2332	1	US-07-864-004B-4	Sequence 4, Appli
7	127	93.4	2332	1	US-08-251-937A-4	Sequence 4, Appli
8	127	93.4	2332	1	US-08-212-133A-2	Sequence 2, Appli
9	127	93.4	2332	1	US-08-270-594A-2	Sequence 2, Appli
10	127	93.4	2332	1	US-08-474-503-2	Sequence 2, Appli
11	127	93.4	2332	3	US-08-670-707A-2	Sequence 2, Appli
12	127	93.4	2332	3	US-09-037-601-2	Sequence 2, Appli
13	127	93.4	2332	3	US-09-324-867-3	Sequence 3, Appli
14	127	93.4	2332	4	US-09-315-179-2	Sequence 2, Appli
15	127	93.4	2332	4	US-09-523-656-2	Sequence 2, Appli
16	127	93.4	2332	5	PCR-US93-03275-4	Sequence 4, Appli
17	127	93.4	2332	5	PCR-US94-13200-2	Sequence 4, Appli
18	127	93.4	2351	1	US-08-121-202-2	Sequence 2, Appli
19	127	93.4	2351	1	US-08-366-851A-2	Sequence 2, Appli
20	127	93.4	2351	6	5171844-2	Patent No. 5171844
21	127	93.4	2351	6	5422260-1	Patent No. 5422260
22	101	74.3	2343	3	US-09-324-867-2	Sequence 2, Appli
23	94	69.1	2304	3	US-09-324-867-4	Sequence 4, Appli
24	94	69.1	2319	1	US-08-212-133A-8	Sequence 8, Appli
25	94	69.1	2319	1	US-08-474-503-6	Sequence 6, Appli
26	94	69.1	2319	2	US-08-670-707A-6	Sequence 6, Appli
27	94	69.1	2319	3	US-09-037-601-6	Sequence 6, Appli


```
TISSUE TYPE: Liver cDNA sequence
US-08-251-937A-4

Query Match
Best Local Similarity 93.4%; Score 127; DB 1; Length 2332;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPGVKVHLKDFPILPGEI 25
II |||||
DB 484 RPLYSRRLPGVKVHLKDFPILPGEI 508

RESULT 8
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Animal factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; APPLICATION NUMBER: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver cDNA sequence
US-08-212-133A-2

Query Match
Best Local Similarity 93.4%; Score 127; DB 1; Length 2332;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPGVKVHLKDFPILPGEI 25
II |||||
DB 484 RPLYSRRLPGVKVHLKDFPILPGEI 508

RESULT 9
US-08-276-594A-2
; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka
; APPLICANT: SUGAWARA, Keishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-276-594A-2

Query Match
Best Local Similarity 93.4%; Score 127; DB 1; Length 2332;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPGVKVHLKDFPILPGEI 25
II |||||
DB 484 RPLYSRRLPGVKVHLKDFPILPGEI 508

RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
```

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,503
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pratt, John S.
;; REGISTRATION NUMBER: 29,476
;; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 404-815-6500
;; TELEFAX: 404-815-6555
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2332 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHEICAL: YES
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapien
;; TISSUE TYPE: Liver CDNA sequence
US-08-474-503-2

Query Match 93.4%; Score 127; DB 1; Length 2332;

Best Local Similarity 96.0%; Pred. No. 7.4e-11;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
Db 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 11
US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver CDNA sequence

;; NAME: Greenlee, Lorraine L.
;; REGISTRATION NUMBER: 27,894
;; REFERENCE/DOCKET NUMBER: 75-95F
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 303/499-8080
;; TELEFAX: 303/499-8089
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2332 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: protein
;; HYPOTHEICAL: YES
;; ANTI-SENSE: NC
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; TISSUE TYPE: Liver
US-08-670-707A-2

Query Match 93.4%; Score 127; DB 2; Length 2332;

Best Local Similarity 96.0%; Pred. No. 7.4e-11;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
Db 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 12

US-09-037-601-2

; Sequence 2, Application US/09037601

; Patent No. 6180371

; GENERAL INFORMATION:

; APPLICANT: Lollar, John S.

; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

; STREET: 5370 Manhattan Circle Suite 201

; CITY: Boulder

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/037,601

; FILING DATE: 26-JUN-1996

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US94/13200

; FILING DATE: 15-NOV-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/212,133

; FILING DATE: 11-MAR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/864,004

; FILING DATE: 07-APR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Ferber, Donna M.

; REGISTRATION NUMBER: 33,878

; REFERENCE/DOCKET NUMBER: 75-95F

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 303/499-8080

; TELEFAX: 303/499-8089

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2332 amino acids

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-09-037-601-2

Query Match          93.4%; Score 127; DB 3; Length 2332;
Best Local Similarity 96.0%; Pred. No. 7.4e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 13
US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
; APPLICANT: Lilliecrap, David
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 6251632ley, Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669.0010002/JAG/BJD
; CURRENT APPLICATION NUMBER: US/09/324.867A
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-324-867-3

Query Match          93.4%; Score 127; DB 3; Length 2332;
Best Local Similarity 96.0%; Pred. No. 7.4e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 485 RPLYSRRLPKGVKHLKDFPILPGEI 509

RESULT 14
US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-09-037-601-2

; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-179-2

Query Match          93.4%; Score 127; DB 4; Length 2332;
Best Local Similarity 96.0%; Pred. No. 7.4e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 15
US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523.656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-656-2

Query Match          93.4%; Score 127; DB 4; Length 2332;
Best Local Similarity 96.0%; Pred. No. 7.4e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

Search completed: October 17, 2003, 11:45:06
Job time : 31 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:36:06 : Search time 69 Seconds
(without alignments)
59.390 Million cell updates/sec

Title: SEQID2-NAT3

Perfect score: 136

Sequence: 1 RPNYSRRLPKGVKHLKDFPILPGEI 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 609560

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	93.4	1438	14	US-10-006-091-1
2	127	93.4	1438	14	US-10-047-257-1
3	127	93.4	1438	15	US-10-225-900-1
4	127	93.4	1471	14	US-10-095-718-2
5	127	93.4	2332	10	US-09-557-641-2
6	127	93.4	2332	12	US-10-131-510A-2
7	127	93.4	2332	15	US-10-187-319-2
8	127	93.4	2351	15	US-10-132-829-4
9	127	93.4	2351	15	US-10-172-712-27
10	101	74.3	1431	14	US-10-095-718-4
11	94	69.1	2319	12	US-10-131-510A-6
12	94	69.1	2319	15	US-10-187-319-6
13	75	55.1	368	12	US-10-131-510A-4
14	75	55.1	368	15	US-10-187-319-4
15	75	55.1	1443	12	US-10-131-510A-59

16	75	55.1	1443	15	US-10-187-319-39	Sequence 39, Appl
17	75	55.1	2133	12	US-10-131-510A-37	Sequence 37, Appl
18	75	55.1	2133	15	US-10-187-319-37	Sequence 37, Appl
19	50	36.8	897	10	US-09-549-192-49	Sequence 49, Appl
20	48.5	35.7	405	10	US-09-863-475A-8	Sequence 8, Appl
21	48.5	35.7	405	12	US-10-234-041-11	Sequence 11, Appl
22	48.5	35.7	432	9	US-09-733-524-7	Sequence 7, Appl
23	48.5	35.7	433	12	US-10-189-977-11	Sequence 11, Appl
24	48.5	35.7	433	12	US-10-392-098-11	Sequence 11, Appl
25	48.5	35.7	433	14	US-10-120-319-11	Sequence 11, Appl
26	48.5	35.7	530	12	US-10-234-041-13	Sequence 13, Appl
27	48	35.3	499	15	US-10-156-761-11906	Sequence 11906, A
28	47	34.6	345	10	US-09-738-626-6584	Sequence 6584, Ap
29	47	34.6	826	11	US-09-988-626-2	Sequence 2, Appl
30	47	34.6	826	11	US-09-988-687-2	Sequence 2, Appl
31	47	34.6	826	11	US-09-988-686-2	Sequence 2, Appl
32	45	33.1	444	15	US-10-157-223-8	Sequence 8, Appl
33	44.5	32.7	331	12	US-10-278-946-18	Sequence 18, Appl
34	44.5	32.7	376	10	US-09-834-722-2	Sequence 2, Appl
35	44.5	32.7	441	10	US-09-738-626-4005	Sequence 4005, Ap
36	44	32.4	282	9	US-09-893-637-2	Sequence 2, Appl
37	44	32.4	282	9	US-09-970-711-21	Sequence 21, Appl
38	44	32.4	388	15	US-10-156-761-14067	Sequence 14067, A
39	44	32.4	482	9	US-09-796-858-40	Sequence 40, Appl
40	44	32.4	491	10	US-09-597-664-3	Sequence 3, Appl
41	44	32.4	826	11	US-09-988-626-226	Sequence 226, App
42	44	32.4	826	11	US-09-988-687-226	Sequence 226, App
43	44	32.4	826	11	US-09-988-686-226	Sequence 226, App
44	44	32.4	913	12	US-09-576-782-94	Sequence 94, Appl
45	44	32.4	1142	12	US-09-576-782-92	Sequence 92, Appl

ALIGNMENTS

RESULT 1
US-10-006-091-1
: Sequence 1, Application US/10006091
: Publication No. US20020102730A1
: GENERAL INFORMATION:
: APPLICANT: Cho, Myung-Sam
: APPLICANT: Chan, Sham-Yuen
: APPLICANT: Kelsey, William
: APPLICANT: Yee, Helena
: TITLE OF INVENTION: Expression System for Factor VIII
: FILE REFERENCE: MSB-7255.1
: CURRENT APPLICATION NUMBER: US/10/006,091
: CURRENT FILING DATE: 2001-12-06
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1438
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Derived from
: OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

Query Match 93.4%; Score 127; DB 14; Length 1438;
Best Local Similarity 96.0%; Pred. No. 1.3e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRRLPKGVKHLKDFPILPGEI 25
|||
DB 484 RPLYSRRLPKGVKHLKDFPILPGEI 508

RESULT 2
US-10-047-257-1
: Sequence 1, Application US/10047257
: Publication No. US20020115152A1
: GENERAL INFORMATION:

TELEFAX: 303/439-8005
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single

```

; SEQ ID NO: 27
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-172-712-27

```


Query Match 93.4%; Score 127; DB 15; Length 2351;
Best Local Similarity 96.0%; Pred. No. 2.3e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPNYSRLPKGVKHLKDFPILPGEI 25
II :|||||:|||||:|||||:|||||
Db 503 RPLYSRLPKGVKHLKDFPILPGEI 527

RESULT 10

US-10-095-718-4

; Sequence 4, Application US/10095718
; Publication No. US20020131956A1

; GENERAL INFORMATION:

; APPLICANT: Walsh, Christopher

; APPLICANT: Chao, Hengjun

; APPLICANT: Burstein, Haim

; APPLICANT: Lynch, Carmel

; APPLICANT: Stepan, Tony

; APPLICANT: Munson, Keith

; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and

; FILE REFERENCE: 35052/204375

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: US/10/095,718

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/158,780

; PRIOR FILING DATE: 1999-10-12

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 1431

; TYPE: PRT

; ORGANISM: canine B-domain deleted factor VIII

US-10-095-718-4

Query Match 74.3%; Score 101; DB 14; Length 1431;

Best Local Similarity 79.2%; Pred. No. 1.3e-06;

Matches 19; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYSRRLPKGVKHLKDFPILPGEI 25

I : : |||||:|||||:|||||:|||||

Db 498 PLHTGRLPKGVKHLKDFPILPGEI 521

RESULT 11

US-10-131-510A-6

; Sequence 6, Application US/10131510A

; Publication No. US20030166536A1

; GENERAL INFORMATION:

; APPLICANT: Lollar, John S

; TITLE OF INVENTION: Modified Factor VIII

; FILE REFERENCE: 75-95J

; CURRENT APPLICATION NUMBER: US/10/131,510A

; CURRENT FILING DATE: 2002-07-10

; PRIOR APPLICATION NUMBER: U.S. 09/315,179

; PRIOR FILING DATE: 1999-05-20

; PRIOR APPLICATION NUMBER: U.S. 09/037,601

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: U.S. 08/670,707

; PRIOR FILING DATE: 1996-06-26

; PRIOR APPLICATION NUMBER: PCT/US97/11155

; PRIOR FILING DATE: 1997-06-26

; PRIOR APPLICATION NUMBER: PCT/US94/13200

; PRIOR FILING DATE: 1994-11-15

; PRIOR APPLICATION NUMBER: U.S. 08/212,133

; PRIOR FILING DATE: 1994-03-11

; PRIOR APPLICATION NUMBER: U.S. 07/864,004

; PRIOR FILING DATE: 1992-04-07

; NUMBER OF SEQ ID NOS: 40

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6
; LENGTH: 2319
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-131-510A-6

Query Match 69.1%; Score 94; DB 12; Length 2319;
Best Local Similarity 66.7%; Pred. No. 2.6e-05;
Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYSRRLPKGVKHLKDFPILPGEI 25

I : : |||||:|||||:|||||:|||||

Db 504 PLHARLPRGIKHVKDLPIHPGEI 527

RESULT 12

US-10-187-319-6

; GENERAL INFORMATION:

; APPLICANT: Lollar, John S.

; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

; STREET: 5370 Manhattan Circle Suite 201

; CITY: Boulder

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/187,319

; FILING DATE: 27-Aug-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/523,656

; FILING DATE: 2000-03-10

; APPLICATION NUMBER: US 09/037,601

; FILING DATE: 1998-03-10

; APPLICATION NUMBER: WO PCT/US97/11155

; FILING DATE: 1997-06-26

; APPLICATION NUMBER: US 08/670,707

; FILING DATE: 1996-05-26

; ATTORNEY/AGENT INFORMATION:

; NAME: Greenlee, Lorraine L.

; REGISTRATION NUMBER: 27,894

; REFERENCE/POCKET NUMBER: 75-95K

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 303/499-8080

; TELEFAX: 303/499-8089

; TITLE: Sequence of the Murine Factor VIII cDNA

; JOURNAL: Genomics

; VOLUME: 16

; PAGES: 374-379

; DATE: 1993

; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-187-319-6

Query Match

Best Local Similarity 66.7%; Score 94; DB 15; Length 2319;

Matches 16; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYSRRLPKGVKHLKDFPILPGEI 25

I : : |||||:|||||:|||||:|||||

Db 504 PLHARLPRGIKHVKDLPIHPGEI 527

RESULT 13

US-10-131-510A-4

```

; Sequence 4, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 06/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-131-510A-4

Query Match 55.1%; Score 75; DB 12; Length 368;
Best Local Similarity 83.3%; Pred. No. 0.0027;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 RLPGVKVHLKDFPILPGE 24
   || || ||||| |||||
Db 118 RLKGWKLKDMPLPGE 135

RESULT 14
US-10-187-319-4
; Sequence 4, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.50
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:

; Sequence 4, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-131-510A-4

Query Match 55.1%; Score 75; DB 15; Length 368;
Best Local Similarity 83.3%; Pred. No. 0.0027;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 RLPGVKVHLKDFPILPGE 24
   || || ||||| |||||
Db 118 RLKGWKLKDMPLPGE 135

RESULT 15
US-10-131-510A-39
; Sequence 39, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 39
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: factor VIII lacking
; OTHER INFORMATION: the B domain
US-10-131-510A-39

```

Query Match 55.1%; Score 75; DP 12; Length 1443;
 Best Local Similarity 83.3%; Pred. No. 0.012;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 7 RLPKGVKHLKDFPILPGE 24
 Db 509 RLLKGWKKLKPILPGE 526

Search completed: October 17, 2003, 11:46:54
 Job time : 70 secs